```
Identities = 78/96 (81%), Positives = 87/96 (90%)

Query: 2   QSIDLLVNNAGLALGLDKSYEADFGDWMTMINTNVVGLLYLTRCILPKMVEVNRGLIINL 61
           Q I +LVNNAGLALGLDK+YEADF +WMTMINTN+VGLLYLTR +LP MV  + G+IINL
Sbjct: 82  QDITILVNNAGLALGLDKAYKADFENWMTMINTNIVGLLYLTRQLLPHMVSKDEGIIINL 141

Query: 62  GSXAGTIPYPGANVYGASKAFVKQFSLNLRADLAGT 97
           GS AGTIPYPGAN+YGASKAFVKQFSLNLRADLAG+
Sbjct: 142 GSTAGTIPYPGANIYGASKAFVKQFSLNLRADLAGS 177
```

Based on this analysis, it was predicted that these proteins and their epitopes could be useful antigens for vaccines or diagnostics.

Example 2339

A DNA sequence (GBSx2492) was identified in *S.agalactiae* <SEQ ID 7171> which encodes the amino acid sequence <SEQ ID 7172>. This protein is predicted to be mercuric reductase. Analysis of this protein sequence reveals the following:

```
Possible site: 53
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2115(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has homology with the following sequences in the GENPEPT database.

```
>GP:CAC14663 GB:Y10855 mercuric reductase [Bacillus licheniformis]
 Identities = 68/104 (65%), Positives = 82/104 (78%)

Query: 1   MNKFKVNISGMTCTGCEKHVESALEKIGAKNIESSYRRGEAVFELPDDIEVESAIKAIDE 60
           M K++VN+ GMTCTGCE+HV  ALE +GAK IE  YRRGEAVFELP+ +EVE+A KAI E
Sbjct: 1   MKKYRVNVQGMTCTGCEEHVAVALENMGAKRIEVDYRRGEAVFELPNGLEVETAKKAIAE 60

Query: 61  ANYQAGEIEEVSSLENVALINEDNYDLLIIGSGAAFSSAIKAI 104
           A YQ GE EEV S E + L +E +YD +IIGSG AAFSSAI+A+
Sbjct: 61  AKYQPGEAEEVQSQELIQLGDEGDYDYIIIGSGGAAFSSAIEAV 104
```

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2340

A DNA sequence (GBSx2494) was identified in *S.agalactiae* <SEQ ID 7173> which encodes the amino acid sequence <SEQ ID 7174>. Analysis of this protein sequence reveals the following:

```
Possible site: 58
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3341(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has no significant homology with any sequences in the GENPEPT database.

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2341

A DNA sequence (GBSx2495) was identified in *S.agalactiae* <SEQ ID 7175> which encodes the amino acid sequence <SEQ ID 7176>. Analysis of this protein sequence reveals the following:

```
Possible site: 31
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4989(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has no significant homology with any sequences in the GENPEPT database.

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2342

A DNA sequence (GBSx2496) was identified in *S.agalactiae* <SEQ ID 7177> which encodes the amino acid sequence <SEQ ID 7178>. Analysis of this protein sequence reveals the following:

```
Possible site: 30
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2569(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Exatuple 2343

A DNA sequence (GBSx2497) was identified in *S.agalactiae* <SEQ ID 7179> which encodes the amino acid sequence <SEQ ID 7180>. This protein is predicted to be DNA polymerase III alpha subunit (dnaE). Analysis of this protein sequence reveals the following:

```
Possible site: 60
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3124(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

A related DNA sequence was identified in *S.pyogenes* <SEQ ID 4095> which encodes the amino acid sequence <SEQ ID 4096>. Analysis of this protein sequence reveals the following:

```
Possible site: 36
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
```

```
bacterial cytoplasm --- Certainty=0.2600(Affirmative) < Succ>
 bacterial membrane --- Certainty=0.0000(Not Clear) < Succ>
  bacterial outside --- Certainty=0.0000(Not Clear) < Succ>
```

An alignment of the GAS and GBS proteins is shown below.

```
 Identities = 186/237 (78%), Positives = 214/237 (89%)

Query: 10  DPVKHNLIFERFLNKERYSMPDIDIDLPDIYRGKFLRYVRNRYGSMHSAQIVTFSTFGAK 69
           DPV+H+L+FERFLN+ERYSMPDIDIDLPDIYR EFLRYVRNRYGS HSAQIVTFSTFG K
Sbjct: 321 DPVQHDLLFERFLNKERYSMPDIDIDLPDIYRSEFLRYVRNRYGSDHSAQIVTFSTFGPK 380

Query: 70  QAIRDVFKRPGASEYELTNITKKIHPRDNLTSVYNRNLAPRQLIDSKIEYQKAYDIAKRI 129
           QAIRDVFKRPG  EYSLTN+TKKI F+D+L +VY ++++FRQ+I+S+ E+QKA+ IAKRI
Sbjct: 381 QAIRDVFKRPGVPEYELTNLTKKIGFKDSLATVYEKSISPRQVINSRTEPQKAFAIAKRI 440

Query: 130 EGNPRQTSIHAAGVVMSDDLLTDHIPLKNGEDMMITQYDASSVEDNGLLKMDFLGLRNLT 189
           EGNPRQTSIHAAG+VMSDD LT+HIPLK+G+DMMITQYDA +VE NGLLKMDFLGLRNLT
Sbjct: 441 EGNPRQTSIHAAGIVMSDDAITNHIPLKSGDDMMITQVDAHAVFANGLLKMDFLGLRNLT 500

Query: 190 FVQKMKEKVDKDYGISIQLETIDLEDKETLKLFAAGQTKGIFQFESGSGAINLLRRIR 246
           FVQKM+EKV KDYG  I +  IDLED +TL LFA G TKGIFQFESGSGAINLLRRIK
Sbjct: 501 FVQKMQEKVAKDYGCQIDITAIDLEDPQTLALFAKGDTKGIFQFESGNGAINLLRRIK 557
```

Based on this analysis, it was predicted that these proteins and their epitopes could be useful antigens for vaccines or diagnostics.

Example 2344

A DNA sequence (GBSx2498) was identified in *S.agalactiae* <SEQ ID 7181> which encodes the amino acid sequence <SEQ ID 7182>. This protein is predicted to be a methylase. Analysis of this protein sequence reveals the following:

```
Possible site: 60
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2121(Affirmative) < Succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < Succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < Succ>
```

The protein has homology with the following sequences in the GENPEPT database.

```
>GP:AAG21729 GB:AF116907 putative methylase [Corynebacterium hoagii]
 Identities = 48/160 (30%), Positives = 85/160 (53%), Gaps = 6/160 (3%)

Query: 97   EPDDSRNGHNDTDLEETDNQIPEEEVVETIPEIPVTDPYPPEDLITDPYPKTARDKVETNI 156
             EP+          +  E  +  + ++E        +P TDF     D+     P  A+ +V  NI
Sbjct: 1236 EPEAPTQPEAASAAETAEPAVEQQEPRAGPQSVPATDPALGTDV--HVPSGAKARVRANI 1293

Query: 157  VAIRLVKNLEVEHRNASPSEQELLAKYVGWGGLANEFPDD---YNPFKSKSRRELKSLVT 213
              A RLV  L+ + R A+  BQ +LA++ GWG +  E PD+   +  +++ ER  L  L+
Sbjct: 1294 AAARLVLELDEQQRPATAEEQAVLAQWSGWGAVP-EVPDNRSKFLSEWADERAALLDLLG 1352

Query: 214  DKEYSDMKQSSLTAYYTDPSLIRQMWGIVERDGPTGWQIL 253
             +K +S  ++++L A+YTDP+++ ++W  V+R G        +L
Sbjct: 1353 EKGPSQARETTLNAHYTDPAIVGELWRAVQRAGLPDGALL 1392
```

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2345

A DNA sequence (GBSx2499) was identified in *S.agalactiae* <SEQ ID 7183> which encodes the amino acid sequence <SEQ ID 7184>. Analysis of this protein sequence reveals the following:

```
Possible site: 34
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1111(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has no significant homology with any sequences in the GENPEPT database.

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2346

A DNA sequence (GBSx2501) was identified in *S.agalactiae* <SEQ ID 7185> which encodes the amino acid sequence <SEQ ID 7186>. Analysis of this protein sequence reveals the following:

```
Possible site: 39
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4752(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has homology with the following sequences in the GENPEPT database.

```
>GP:CAA61516 GB:X89232 DNA-directed RNA polymerase [Pediococcus
          acidilactici]
 Identities = 48/53 (90%), Positives = 52/53 (97%)

Query: 5  KKPETINYRTLKPEREGLFDEVIFGPTKDWECACGKYKRIRYKGIICDRCGVE 57
          KKPETINYRTLKPE++GLFDE IFGPTKD+ECACGKYKRIRYKGI+CDRCGVE
Sbjct: 29 KKPETINYRTLKPEKDGLFDERIFGPTKDYECACGKYKRIRYKGIVCDRCGVE 81
```

There is also homology to SEQ ID 384.

Based on this analysis, it was predicted that these proteins and their epitopes could be useful antigens for vaccines or diagnostics.

Example 2347

A DNA sequence (GBSx2502) was identified in *S.agalactiae* <SEQ ID 7187> which encodes the amino acid sequence <SEQ ID 7188>. Analysis of this protein sequence reveals the following:

```
Possible site: 22
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3080(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has homology with the following sequences in the GENPEPT database.

```
>GP:AAC00282 GB:AP000220 YtlR [Bacillus subtilis]
 Identities = 61/216 (28%), Positives = 98/216 (45%), Gaps = 28/216 (12%)

Query: 8   IPCTYYPVGSGNDFARALKIPNL---------KETLYAIQTERLKEINCFIYDKGLIL-- 56
            I  ++ P G+ NDF+R   I  +          K  LT  +T  L  +N F+ DK  IL
Sbjct: 86  IELSFVPAGAYNDPSRGPSIKKIDLIQEIKKVKRPLT--RTPHLGSVN-PLQDKSQILYF 142

Query: 57  -NSLDLGPAAYVVWKASNSKCKNILNRYRLGKITYIVIAIKSLLHSSK------VQVLVE 109
             N + +GP AYV  KA    ++ +    RL  + Y +    S LH+S        +    E
Sbjct: 143 MNHIGIGPDAYVNKKAMEPPLRRVLPLRLRPLVYPL----SHLEASATPKDPTLACTTE 198

Query: 110 GETGQQIKLNDLYPFALANNTYPGGGITTWPKASALTAELDMVYAKGHTFLKRLSILLSL 169
              ET +    +D++F  ++N+ ++GGG+   P A+       D+V  +   FLK+  +L  +
Sbjct: 199 DETRE---FHDVWFAVVSNHPFYGGGMKAAPLANPREKTFDIVIVENQPFLKKYWLLCLM 255

Query: 170 VFKRHTTSKSIKHQTFKAMTVYFPKNSLIEIDGEIV 205
             F +HT    +    K +T Y          DGEI+
Sbjct: 256 APGKHTKMDGVTMFKAKDITFYTKDKIPFHADGEIM 291
```

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2348

A DNA sequence (GBSx2503) was identified in *S.agalactiae* <SEQ ID 7189> which encodes the amino acid sequence <SEQ ID 7190>. This protein is predicted to be protease subunit HflC (hflC). Analysis of this protein sequence reveals the following:

```
Possible site: 18
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.1809(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has homology with the following sequences in the GENPEPT database.

```
>GP:AAG08326 GB:AE004907 protease subunit HflC [Pseudomonas aeruginosa]
 Identities = 182/202 (90%), Positives = 194/202 (95%)

Query: 1   MSQTERAVLLQFGKVVQTDVKPGLHVKVPYVNQVRKFDGRLLTLDAPTQRFLTLEKKAVM 60
            + QTERAV+L+FG+VV++DVKPGLH K+PYVNQVRKFD RLLTLDAPTQRFLTLEKKAVM
Sbjct: 26  VQQTERAVMLRFGRVVESDVKPGLHFKIPYVNQVRKFDARLLTLDAPTQRFLTLEKKAVM 85

Query: 61  VDAYAKWRVKDAERFYTATSGLKQIADERLSRRLESGLRDQFGKRTLHEVVSGERDALMA 120
            VDAYAKWRV DAERFYTATSGLKQIADERLSRRLE+GLRDQFGKRTLHEVVSGERDALM
Sbjct: 86  VDAYAKWRVADAERFYTATSGLKQIADERLSRRLEAGLRDQFGKRTLHEVVSGERDALMG 145

Query: 121 DITGSLNRMAKKLGIRVLDVRVKAIDLPKEVNRSVFERMSTERERRAREHRAKGNELGE 180
            DIT SLNRMA+KKLGIEV+DVRVKAIDLPKEVNRSVFERMSTERERRAREHRAKG EL E
Sbjct: 146 DITASLNRMAQKKLGIEVIDVRVKAIDLPKEVNRSVFERMSTERERRAREHRAKGRELAE 205

Query: 181 GIRADADRQRRVLLAEAYRESE 202
            GIRADADRQRRV++AEAYRESE
Sbjct: 206 GIRADADRQRRVIVAEAYRESE 227
```

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2349

A DNA sequence (GBSx2504) was identified in *S.agalactiae* <SEQ ID 7191> which encodes the amino acid sequence <SEQ ID 7192>. Analysis of this protein sequence reveals the following:

```
Possible site: 24
>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

The protein has no significant homology with any sequences in the GENPEPT database.

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2350

A DNA sequence (GBSx2505) was identified in *S.agalactiae* <SEQ ID 7193> which encodes the amino acid sequence <SEQ ID 7194>. This protein is predicted to be ABC transporter (ATP-binding; daunorubicin resistance). Analysis of this protein sequence reveals the following:

```
Possible site: 56
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.1846(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has homology with the following sequences in the GENPEPT database.

```
>GP:CAB15892 GB:Z99123 similar to ABC transporter (ATP-binding
        protein) [Bacillus subtilis]
 Identities = 88/231 (38%), Positives = 132/231 (57%), Gaps = 13/231 (5%)

Query: 10  QVIGYLPDVPKFYDYMTAQEYLQLC---AGLAQNKTSLPIADLLEQVGLADN-QQRISTY 65
           ++IGYLP  P FY +MTA E+L       +GL++ K    I ++LE VGL +   +RI  Y
Sbjct: 69  RLIGYLPQYPAFYSWMTANEFLTFAGRLSGLSKRKCQSKIGEMLERVGLHEAAHKRIGGY 128

Query: 66  SRGMKQRLGLAQALIHXXKILICDEPTSALDPQGRQEILSIISQLRGQKTVIFSTHILSD 125
           S GMKQRLGLAQAL+H  K LI DEP SALDP GR E+L ++ +L+     V+FSTH+L D
Sbjct: 129 SGGMKQRLGLAQALLHKPKFLILDEPVSALDPTGRFEVLDMMRELKKHMAVLFSTHVLHD 188

Query: 126 VEKVCDQVLILLKSGIH---NLEDLRDKASASVNQLNLLIKVSDNSAQKLALRFPLNQKD 182
            E+VCDQV+I+     I     L++L+ +   +V  L++  K+     +K  +   + +
Sbjct: 189 AEQVCDQVVIMKNGEISWKGELQELKQQQQTNVPTLSVKEKLEGWLEEKPYVSAIVYKNP 248

Query: 183 QYYKVHLRLSEANNREQALASFYRYLVEQETTPYFTKLLRDSLRDFYLEVI 233
                +   EL + +      L+       + + +T    E   +SLED YL+V+
Sbjct: 249 S--QAVFELPDIHAGRSLLSD----CIRKGLTVTRFEQKTESLEDVYLKVV 293
```

There is also homology to SEQ ID 686.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2351

A DNA sequence (GBSx2506) was identified in *S.agalactiae* <SEQ ID 7195> which encodes the amino acid sequence <SEQ ID 7196>. Analysis of this protein sequence reveals the following:

```
Possible site: 52
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0679(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has homology with glycine-rich cell wall proteins (*e.g.* GB:AL161589 – the glycine-rich cell wall protein from *Arabidopsis thaliania*) and to SEQ ID 6882.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2352

A DNA sequence (GBSx2507) was identified in *S.agalactiae* <SEQ ID 7197> which encodes the amino acid sequence <SEQ ID 7198>. Analysis of this protein sequence reveals the following:

```
Possible site: 35
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2890(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has no significant homology with any sequences in the GENPEPT database.

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2353

A DNA sequence (GBSx2508) was identified in *S.agalactiae* <SEQ ID 7199> which encodes the amino acid sequence <SEQ ID 7200>. Analysis of this protein sequence reveals the following:

```
Possible site: 60
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2410(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

A related GBS nucleic acid sequence <SEQ ID 9329> which encodes amino acid sequence <SEQ ID 9330> was also identified.

The protein has no significant homology with any sequences in the GENPEPT database.

No corresponding DNA sequence was identified in *S.pyogenes*.

SEQ ID 9330 (GBS678) was expressed in *E.coli* as a His-fusion product. SDS-PAGE analysis of total cell extract is shown in Figure 163 (lane 18; MW 53kDa), Figure 164 (lane 2 & 3; MW 53kDa) and Figure 188 (lane 7; MW 53kDa). Purified protein is shown in Figure 242, lanes 6 & 7.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2354

A DNA sequence (GBSx2509) was identified in *S.agalactiae* <SEQ ID 7201> which encodes the amino acid sequence <SEQ ID 7202>. This protein is predicted to be surface protein Rib. Analysis of this protein sequence reveals the following:

```
Possible site: 24
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2025(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2355

A DNA sequence (GBSx2510) was identified in *S.agalactiae* <SEQ ID 7203> which encodes the amino acid sequence <SEQ ID 7204>. This protein is predicted to be surface protein Rib. Analysis of this protein sequence reveals the following:

```
Possible site: 24
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1892(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2356

A DNA sequence (GBSx2511) was identified in *S.agalactiae* <SEQ ID 7205> which encodes the amino acid sequence <SEQ ID 7206>. This protein is predicted to be surface protein Rib. Analysis of this protein sequence reveals the following:

```
Possible site: 24
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1892(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2357

A DNA sequence (GBSx2512) was identified in *S.agalactiae* <SEQ ID 7207> which encodes the amino acid sequence <SEQ ID 7208>. Analysis of this protein sequence reveals the following:

```
Possible site: 28
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0999(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has no significant homology with any sequences in the GENPEPT database.

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2358

A DNA sequence (GBSx2514) was identified in *S.agalactiae* <SEQ ID 7209> which encodes the amino acid sequence <SEQ ID 7210>. This protein is predicted to be surface protein Rib. Analysis of this protein sequence reveals the following:

```
Possible site: 24
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1892(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2359

A DNA sequence (GBSx2515) was identified in *S.agalactiae* <SEQ ID 7211> which encodes the amino acid sequence <SEQ ID 7212>. Analysis of this protein sequence reveals the following:

```
Possible site: 19
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2041(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

The protein has no significant homology with any sequences in the GENPEPT database.

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2360

A DNA sequence (GBSx2516) was identified in *S.agalactiae* <SEQ ID 7213> which encodes the amino acid sequence <SEQ ID 7214>. This protein is predicted to be 30S ribosomal protein S6 (rpsF). Analysis of this protein sequence reveals the following:

```
Possible site: 51
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3607(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

A related GBS nucleic acid sequence <SEQ ID 9423> which encodes amino acid sequence <SEQ ID 9424> was also identified.

The protein has homology with the following sequences in the GENPEPT database.

```
>GP:CAB16128 GB:Z99124 ribosomal protein S6 (BS9) [Bacillus subtilis]
 Identities = 41/72 (56%), Positives = 58/72 (79%), Gaps = 1/72 (1%)

Query: 1  MVARFDSILSDNGATVVESKDWEKRRLAYSIQDFTEGLYHIVNVEAEDAVALNEFDRLSK 60
          ++ RF+++L+ NGA +  +KDW KRRLAYEI DF +G Y IVNV++ DA A+ EFDRL+K
Sbjct: 22 VIRRFNNVLTSNGAEITGTKDWGKRRLAYEINDFRDGFYQIVNVQS-DAAAVQEFDRLAK 80

Query: 61 INGDILRHMIVK 72
          I+ DI+RH++VK
Sbjct: 81 ISDDIIRHIVVK 92
```

A related DNA sequence was identified in *S.pyogenes* <SEQ ID 7215> which encodes the amino acid sequence <SEQ ID 7216>. Analysis of this protein sequence reveals the following:

```
Possible site: 40
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2720(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

An alignment of the GAS and GBS proteins is shown below.

```
 Identities = 66/74 (89%), Positives = 70/74 (94%)

Query: 1  MVARFDSILSDNGATVVESKDWEKRRLAYKIQDFTEGLYHIVNVEAEDAVALNEFDRLSK 60
          +VARFDSIL+DNGATVVESKDWEKRRLAYEI DF EGLYHIVN+EA DA ALNEFDRLSK
Sbjct: 22 LVARFDSILTDNGATVVESKDWEKRRLAYEINDFREGLYHIVNLEATDAAALNEFDRLSK 81

Query: 61 INGDILRHMIVKVD 74
          INGDILRHMIVK+D
Sbjct: 82 INGDILRHMIVKLD 95
```

Based on this analysis, it was predicted that these proteins and their epitopes could be useful antigens for vaccines or diagnostics.

Example 2361

A DNA sequence (GBSx2518) was identified in *S.agalactiae* <SEQ ID 7219> which encodes the amino acid sequence <SEQ ID 7220>. This protein is predicted to be surface protein Rib. Analysis of this protein sequence reveals the following:

```
Possible site: 49
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.5289(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.pyogenes*.

Based on this analysis, it was predicted that this protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2362

A DNA sequence (GASx1R) was identified in *S.pyogenes* <SEQ ID 7221> which encodes the amino acid sequence <SEQ ID 7222>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2363

A DNA sequence (GASx5R) was identified in *S.pyogenes* <SEQ ID 7223> which encodes the amino acid sequence <SEQ ID 7224>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2364

A DNA sequence (GASx11) was identified in *S.pyogenes* <SEQ ID 7225> which encodes the amino acid sequence <SEQ ID 7226>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2614(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2365

A DNA sequence (GASx17) was identified in *S.pyogenes* <SEQ ID 7227> which encodes the amino acid sequence <SEQ ID 7228>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2849(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2366

A DNA sequence (GASx18) was identified in *S.pyogenes* <SEQ ID 7229> which encodes the amino acid sequence <SEQ ID 7230>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2099(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2367

A DNA sequence (GASx34) was identified in *S.pyogenes* <SEQ ID 7231> which encodes the amino acid sequence <SEQ ID 7232>. Analysis of this protein sequence reveals the following:

```
Possible site: 54

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0801(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2368

A DNA sequence (GASx38) was identified in *S.pyogenes* <SEQ ID 7233> which encodes the amino acid sequence <SEQ ID 7234>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB12617 GB:Z99108 similar to protein-tyrosine phosphatase
             [Bacillus subtilis]
 Identities = 57/155 (36%), Positives = 88/155 (56%), Gaps = 12/155 (7%)

Query: 1   MKKVCFVCLGNICRSPMAEFVMKSIVS----SDVMMIESRATSDWEHGNDIHSGTQSILK 56
            M  V FVCLGNICRSPMAE + + + +        +  +S     W  GNP H GTQ IL+
Sbjct: 1   MISVLFVCLGNICRSPMAHAIFRDLAAKKGLSGKIKADSAGIGGWHIGNPPHEGTQEILR 60

Query: 57  TYQINYDITKCSKQITTTDFNTFDYIIGMDSDNVKNLKEMSQKQWDSKIYLFKE------ 110
               I++D    ++Q++  D + EDYII MD++N+ +L+ M+  +  S I    +
Sbjct: 61  REGISFD-GMLARQVSEQDLEDFDYIIAMDANNIGSLRSMAGFKNTSHIKRLLDYVEDSD 119

Query: 111 -GSVPDPWYTMDFEETYQLVRKGCQDWLSRLMSKE 144
             VPDP+YT +FEE  QL++ GC+  I+ +  ++
Sbjct: 120 LADVPDPYYTGNFEEVCQLIKTGCEQLLASIQKEK 154
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2369

A DNA sequence (GASx42R) was identified in *S.pyogenes* <SEQ ID 7235> which encodes the amino acid sequence <SEQ ID 7236>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4753(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2370

A DNA sequence (GASx47R) was identified in *S.pyogenes* <SEQ ID 7237> which encodes the amino acid sequence <SEQ ID 7238>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2014(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2371

A DNA sequence (GASx53R) was identified in *S.pyogenes* <SEQ ID 7239> which encodes the amino acid sequence <SEQ ID 7240>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.11   Transmembrane   56 -  72 (  56 -  72)

----- Final Results -----
             bacterial membrane --- Certainty=0.1044(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2372

A DNA sequence (GASx67R) was identified in *S.pyogenes* <SEQ ID 7241> which encodes the amino acid sequence <SEQ ID 7242>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1610(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2373

A DNA sequence (GASx75) was identified in *S.pyogenes* <SEQ ID 7243> which encodes the amino acid sequence <SEQ ID 7244>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2803(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA41942 GB:X59250 ribosomal protein B [Lactococcus lactis]
 Identities = 37/38 (97%), Positives = 37/38 (97%)

Query: 1  MKVRPSVKPICKYCKVIRRNGRVMVICPTNPKHKQRQG 38
          MKVRPSVKPICEYCKVIRRNGRVMVICP NPKHKQRQG
Sbjct: 1  MKVRPSVKPICEYCKVIRRNGRVMVICPANPKHKQRQG 38
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2374

A DNA sequence (GASx76) was identified in *S.pyogenes* <SEQ ID 7245> which encodes the amino acid sequence <SEQ ID 7246>. Analysis of this protein sequence reveals the following:

```
Possible site: 35

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
```

```
bacterial cytoplasm --- Certainty=0.0824(Affirmative) < succ>
 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB06824 GB:L47971 ribosomal protein S13 [Bacillus subtilis]
 Identities = 86/121 (71%), Positives = 103/121 (85%)

Query: 1   MARIAGVDIPNDKRVVISLTYVYGIGLATSKKILAAAGISEDIRVKDLTSDQEDAIRREV 60
            MARIAGVDIP DKRVVISLTY++GIG  T++++L  AG+SED RV+DLT ++   IR  +
Sbjct: 1   MARIAGVDIPRDKRVVISLTYIFGIGRTTAQQVLKEAGVSEDIRVRDLTEEELGKIRDII 60

Query: 61  DAIKVEGDLRREVNMNIKRLMEIGSYRGIRHRRGLPVRGQNTKNNARTRKGKAVAIAGKKK 121
            D +KVEGDLRREV++NIKRL+EIGSYRGIRHRRGLPVRGQN+KNNARTRKG     +A KKK
Sbjct: 61  DKLKVEGDLRREVSLNIKRLIEIGSYRGIRHRRGLPVRGQNSKNNARTRKGPRRTVANKKK 121
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2375

A DNA sequence (GASx81R) was identified in *S.pyogenes* <SEQ ID 7247> which encodes the amino acid sequence <SEQ ID 7248>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1842(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2376

A DNA sequence (GASx82) was identified in *S.pyogenes* <SEQ ID 7249> which encodes the amino acid sequence <SEQ ID 7250>. Analysis of this protein sequence reveals the following:

```
Possible site: 59

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3613(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2377

A DNA sequence (GASx83) was identified in *S.pyogenes* <SEQ ID 7251> which encodes the amino acid sequence <SEQ ID 7252>. Analysis of this protein sequence reveals the following:

```
Possible site: 51

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1141(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2378

A DNA sequence (GASx85) was identified in *S.pyogenes* <SEQ ID 7253> which encodes the amino acid sequence <SEQ ID 7254>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2280(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2379

A DNA sequence (GASx89R) was identified in *S.pyogenes* <SEQ ID 7255> which encodes the amino acid sequence <SEQ ID 7256>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3040(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2380

A DNA sequence (GASx102) was identified in *S.pyogenes* <SEQ ID 7257> which encodes the amino acid sequence <SEQ ID 7258>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-13.75   Transmembrane   21 -  37 (  12 -  41)

----- Final Results -----
                bacterial membrane --- Certainty=0.6498(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC45312 GB:U81957 ComYC [Streptococcus gordonii]
 Identities = 59/104 (56%), Positives = 85/104 (81%), Gaps = 1/104 (0%)

Query: 6   NNLRHKKLKGFTLLEMLLVILVISVLMLLFVPNLSKQKDRVTETGNAAVVKLVENQAELY 65
           N L+  ++K FTL+EML+V+L+ISVLMLLFVPNL+KQK+ V++TGNAAVVK+VE+QAELY
Sbjct: 2   NKLKKLRVKAFTLVEMLVVLLIISVLMLLFVPNLTKQKEAVSDTGNAAVVKVVESQAELY 61

Query: 66  SL-SQGSKPSLSQLKADSITEKQEKAYQDYYDKHNEKARLSN 108
           SL + G + +LS+L A G+I++KQ  +Y+ YY K+ +E   ++N
Sbjct: 62  SLNTGDQATLSKLVAAGNLSQKQADSYKAYYGKHNSETQAVAN 105
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2381

A DNA sequence (GASx103) was identified in *S.pyogenes* <SEQ ID 7259> which encodes the amino acid sequence <SEQ ID 7260>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC23740 GB:AF052207 competence protein [Streptococcus pneumoniae]
 Identities = 52/131 (39%), Positives = 76/131 (57%)

Query: 8   IKAFTLLETLLSLSVMSFIILGLSVPVTKSYQKVEEHLFFSHPEHLYRHQQKLALLQQKQ 67
           IKAFT+LE+LL L ++S + LGLS  V  ++  VEE +FF  PE LYR  QK ++  Q++
Sbjct: 2   IKAFTMLESLLVLSLVSILALSLSGSVQSTFSAVEEQIFFMEPEELYRETQKRSVASQQK 61
```

```
Query: 68  RVLDISSTKIVTEGNSLTVPKSITVNHPYRLVIDQMGGNHSLAKIIFDMTDRRFKYQFYL 127
             L++    I       LTVPK I      +  D+ GGN SLAK+ F  +     +YQ YL
Sbjct: 62  TSLNLDGQMISNGSQKLTVPKGIQAPSGQSITFDRAGGNSSLAKVEFQTSKGAIRYQLYL 121

Query: 128 GSGNYQKTSQS 138
           G+G  ++  ++
Sbjct: 122 GNGKIKRIKET 132
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2382

A DNA sequence (GASx104) was identified in *S.pyogenes* <SEQ ID 7261> which encodes the amino acid sequence <SEQ ID 7262>. Analysis of this protein sequence reveals the following:

```
Possible site: 23

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2383

A DNA sequence (GASx109) was identified in *S.pyogenes* <SEQ ID 7265> which encodes the amino acid sequence <SEQ ID 7266>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-10.51   Transmembrane   37 -  53 (  28 -  58)
   INTEGRAL    Likelihood = -3.56   Transmembrane   61 -  77 (  60 -  77)

----- Final Results -----
              bacterial membrane --- Certainty=0.5203(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2384

A DNA sequence (GASx115R) was identified in *S.pyogenes* <SEQ ID 7267> which encodes the amino acid sequence <SEQ ID 7268>. Analysis of this protein sequence reveals the following:

```
Possible site: 18
```

```
>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-11.09   Transmembrane   20 -  36 (  13 -  40)

----- Final Results -----
                bacterial membrane --- Certainty=0.5437(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2385

A DNA sequence (GASx124) was identified in *S.pyogenes* <SEQ ID 7269> which encodes the amino acid sequence <SEQ ID 7270>. Analysis of this protein sequence reveals the following:

```
Possible site: 52

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -8.17   Transmembrane   31 -  47 (  29 -  59)
   INTEGRAL    Likelihood = -5.63   Transmembrane  737 - 753 ( 734 - 756)

----- Final Results -----
                bacterial membrane --- Certainty=0.4270(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC97148 GB:U49397 Cpa [Streptococcus pyogenes]
 Identities = 401/737 (54%), Positives = 517/737 (69%), Gaps = 25/737 (3%)

Query: 25  SKNSKR--FTVTLVGVFLMIFALVTSMVGAKTVFGLVESSTPNAINPDSSSEYRWYGYES 82
           S N+KR    T+ L+ VFL    AL+  +  +    PG  E S PN      S  +Y WYGY+S
Sbjct: 11  SANNKRQVTTIGLLKVFLITFVALIGIVGFSIRAPGABEQSVFN--RQSSIQDYPWYGYDS 68

Query: 83  YVRGHPYYKQFRVAHDLRVNLEGSRSYQVYCPNLKKAPFLGSDSSVKKWYKKHDGISTKF 142
           Y +G+P Y   +  H+L+VNLEGS+ YQ YCPNL K FP  SDS    +WYKK +G +  F
Sbjct: 69  YPKGYPDYSPLKTYHNLKVNLEGSKDYQAYCPNLTKHFPSKSDSVRSQWYKKLBGTNENF 128

Query: 143 EDYAMSPRITGDELNQKLRAVMYNCHPQNANGIMKGLEPLNAIRVTQEAVWYYSDNAPIS 202
             A  PRI    +L Q +  ++YNG+P N NGIM+G++PLNAI VTQ A+W Y+D+A I
Sbjct: 129 IKLADKPRIEDGQLQQNILRILYNGYPNNRNGIMKGIDPLNAILVTQNAIW-YTDSAQI- 186

Query: 203 NPDESPKRESESNLVSTSQLSLMRQALKQLIDPNLATKMPKQVPDDPQLSIPESHDKGDK 262
           NPDESPK E+ SN ++  QL LMR+ALK+LIDPNL +K   + P  ++L++PES D
Sbjct: 187 NPDESFKTEARSNGINDQQLGLMRKALKELIDPNLGSKYSNKTPSGYRLNVPESHD---- 242

Query: 263 YNKGYQNLLSGGLVPTKPFTPGDPPMPPNQPQTTSVLIRKYAIGDYSKLLEGATLQLTGD 322
             K +QNLLS   VP  PP PG+   PP + + TSV+IRKYA GD  SKLLEGATL+L+
Sbjct: 243 --KPFQNLLSAEYVPDTPPKPGEE--PPAKTEKTSVIIRKYAEGD--SKLLEGATLKLSQI 297

Query: 323 NVNSFQARVFSSNDIGERIKLSDGIYTYLTELNSPAGYSIAEPITFKVEAGKVYTI-IDGK 381
           + FQ + F SN +GE +EL +GTYTLTE +SP GY IAEPI F+VE  KV+ +  DG
Sbjct: 298 EGSGPQEKDPQSNSLGSTVELPNGIYTYLTETSSPDGYKIAEPIKFRVENKKVFIVQKDGS 357

Query: 382 QIENPNKEIVEPYSVEAYNDFEEFSVLT-TQNYAKPYYAKMKNGSSQVVYCFNADLKSPP 440
           Q+ENPNKE+ EPYSVEAYNDF +  VL+     Y KPYYA NK+ SSQVVYCFNADL SPP
```

```
Sbjct: 358 QVENPNKEVAEPYSVEAYNDPMDEEVLSGFTPYGKFYYATNKDKSSQVVYCFNADLHSPF 417

Query: 441 DSEDSGKTMTPDFTT-GEVKYTHIAGRDLFKVTVKPRDTDPDTFLKHIKKVIEKGYREKG 499
           DS D G+T+ PD +T  EVKYTH AG DLFKV ++PRDT+P+ FLKHIKKVIEKGY++KG
Sbjct: 418 DSYDSGETINPDTSTMKEVKYTHTAGSDLFKYALRPRDTNPEDFLKHIKKVIEKGYKKKG 477

Query: 500 QAIEYSGLTETQLRAATQLAIYYPTDSAELDKDKL----KDYHGPGDMNDSTLAVAKILV 555
            + Y+GLTETQ RAATQLAIYYPTDSA+L    K     K YHGF  M++ TLAV K L+
Sbjct: 478 DS--YNGLTETQFRAATQLAIYYPTDSADLKTLKTYNNGKGYHGFRSMDEKTLAVTKELI 535

Query: 556 EYAQDSNPPQLTDLDFFIPNNNKYQSLIGTQWHPEDLVDIIRMEDKK-EVIPVTHNLTLR 614
           YAQ+ + PQLT+LDFF+PNN+K QSLIGT+ HP+DLVD+IRMEDKK EVIPVTH+LT++
Sbjct: 536 TYAQNGSAPQLTNLDFFVPNNSKDQSLIGTECHPDDLVDVIRMEDKKQEVIPVTHSLTVK 595

Query: 615 KTVTGLAGDKTKDFHFEIKLKNNKQELLSQTVKTDKTNLEFKDGKATTNLKHGESLTLQG 674
           KTV G  GD+TK F FE+ELK+   + +  T+KT+  +L  KDGK + NLKHG+++ ++G
Sbjct: 596 KTVVGELGDKTKGFQFELELKDKTGQPIVMTLKTNNQDLVAKDGKYSFNLKHGDTIRIBG 655

Query: 675 LPEGYSYLVKETDSEGYKVKVNSQEVANATVSKTGITSDETLAPENNKEPVVPTGVDQKI 734
           LP GYSY +KE +++ Y V V+++     A       IT D+ + PEN K+ V PTG+
Sbjct: 656 LPTGYSYTLKEAEAKDYIVTVDNKVSQEAQSVGKDITEDKKVTPENNRKDLVPPTGLTTDG 715

Query: 735 NGYLALIVIAGISLGIW 751
             YL L+++  + L +W
Sbjct: 716 AIYLWLLLLVPLGLLVW 732
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2386

A DNA sequence (GASx125R) was identified in *S.pyogenes* <SEQ ID 7271> which encodes the amino acid sequence <SEQ ID 7272>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2604(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2387

A DNA sequence (GASx126) was identified in *S.pyogenes* <SEQ ID 7273> which encodes the amino acid sequence <SEQ ID 7274>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1537(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC97149 GB:U49397 LepA [Streptococcus pyogenes]
 Identities = 59/132 (44%), Positives = 84/132 (62%), Gaps = 5/132 (3%)

Query: 1   MIIKRNDMADSVKAGDAILFYRLSQTKVDERAVVYEDSKTSITKVGRIIAQAGDEVDLTE 60
           MII  NDM+D++ AGD +L+YRL+    + + VVYE   T   KVGRI AQAGDEV+ T+
Sbjct: 42  MIINTNDMSPALSAGDGVLYYRLADRSHINDVVYEVDNT--LKVGRIAAQAGDEVNFTQ 99

Query: 61  QGELKINGHIQNBG---LTFIKSRSANYPYRIADNSYLILNDYYSQSESNNYLQDAIAKDA 117
           +G L INGH    +   LT+  S   N+PY++    +Y ILNDY  +  ++     A+  +
Sbjct: 100 EGGLLINGHPQSKEVPYLTYPHSSGPNFPYKVPTGTYFILNDYREKRLDSRYYGLAPINQ 159

Query: 118 IKGTINTLIRLR 129
           IKG I+TL+R+R
Sbjct: 160 IKGKISTLLRVR 171
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2388

A DNA sequence (GASx127) was identified in *S.pyogenes* <SEQ ID 7275> which encodes the amino acid sequence <SEQ ID 7276>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -3.93   Transmembrane  312 - 328 ( 311 - 337)

----- Final Results -----
                bacterial membrane --- Certainty=0.2572(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC97152 GB:U49397 unknown [Streptococcus pyogenes]
 Identities = 125/355 (35%), Positives = 191/355 (53%), Gaps = 26/355 (7%)

Query: 1   MKLRHLLLLTGAAL/TSFA-----ATIVHGET--VVNGAKLTVTKNL-DLVNSNALIPVTDF 52
           MK   LLL  A L +          +  +  ET  V++G+ L V K          + N L+P  D+
Sbjct: 1   MKKNNKLLLATAILATALGMASMSQNIKATAGVIDGSTLVVKKTFPSYTDDNVLMPKADY 60

Query: 53  TFKIEPDTTVN---EDGNKFK-GVALNTPMTK-VTVTNSDKGGSNTKTAREDFSEVTPEK 107
           +FK+E D        +DG   K GV       TK + Y+NSDK  +  K+  F+P+ V F
Sbjct: 61  SFKVEAADDNAKGKTKDGLDIKPGVIDGLENTKTIRYSNSDKITAKEKSVNFEFANVKPG 120

Query: 108 PGVYYYKVTZEKIDKVPGVSYDTTSYTVQVHVLNNBEQQKPVATYIVGYKEGS--KVPIQ 165
            GVY Y V E   +K  G++YD+  +TV V+V+ N+E        YIV  + G    K P+
Sbjct: 121 VGVYRYTVAEVNGNKA-GITYDSQQWTVDVYVV-NKEGGGPEVVKYIVSTEVGQSEKKPVL 178

Query: 166 FKNSLDSTTLIVKKKVSGTGGDRSKDFNFGLTLKANQYYKASEKVMIEKTTKGGQAPVQT 225
           FKNS D+T+L ++K+V+G  G+  + F+F L L  N+ +    EK  +  +GG+
Sbjct: 179 FKNSFDTTSLKIEKQVTGNTGEHQRLFSFTLLLTPNECF---EKGQVVNILQGGETK--- 232

Query: 226 EASIDQLYHFTLKDGESIKVTNLPVGVDYVVTEDDYKSEKYTTNVEVSPQDGAVKNIAGN 285
           +  I + Y FTLKD  S+ ++ LPVG++YV +TE+D   + Y T+  +   + +     G
Sbjct: 233 KVVIGEEVSFTLKDKGSVTLSQLPVGIEYKLTEDDVTKDGYKTSATLKDQESSTTELGK 292

Query: 286 STRQETTSTDKDWTITTFTNKKDFEVPTGVAMTVAPYIALGIVAVGGALYFVKKKNA 340
            + + S D+   I  TNK+D +VPTGV  T+AP+  L IVA+GG +Y  K+K A
```

```
Sbjct: 293 DHKTDKSADE---IVVTNKRDTQVPTGVVGTLAPPAVLSIVAIGGVIYITKRKKA 344
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2389

A DNA sequence (GASx128) was identified in *S.pyogenes* <SEQ ID 7277> which encodes the amino acid sequence <SEQ ID 7278>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC97152 GB:U49397 unknown [Streptococcus pyogenes]
 Identities = 115/240 (47%), Positives = 178/240 (73%), Gaps = 3/240 (1%)

Query: 1   MIVRLIKLLDKLINVIVLCFFFLCLLIAALGIYDALTVYQGANNTNYQQYIKKKGVQ--FD 58
           M++ ++++++K I+ ++L F  + L +A  G++D+  +YQ A+A+N+++++K    Q  F+
Sbjct: 351 MWMTIVQVINKAIDTLILIFCLVVLFLAGFGLWDSYHLYQQADASNFKKFKTAQQQPKFE 410

Query: 59  DLLAINSDVWAWLTVKGTHIDYFIVQGENNLEYINKSVEGEYSLSGSVFLDYRNKVTFED 118
           DLLA+N DV+ WL + GTHIDYF+VQG+ NLEYINK+V+G  ++SGS+FLD RN   F D
Sbjct: 411 DLLALNEDVIGWLNIPGTHIDYFLVQGKTNLEYINKAVDGSVAMSGSLFLDTRNHNDFTD 470

Query: 119 KYSLIYAHHMAGNVMPGRLPNFKKKSPFNKHKEFSIETKTKQKLKINIFACIQTDAFDSL 178
            YSLIY HHMAGN MPGE+P F KK+PFNKH +  IETK ++KL + IFAC++TDAFD L
Sbjct: 471 DYSLIYGHHMAGNAMPGEIPKFLKKNPFNKHNKAIIETKERKKLTVTIFACLKTDAFDQL 530

Query: 179 LFNPIDV-DISSKNEFLNHIKQKSVQYRELITNESRFVALSTCEDMTTDGRIVIGQIE 237
           +FNP  + +   + + +++I ++S Q++ +   + ++FVA STCE+ +TD R+IV+G I+
Sbjct: 531 VFNPNATTNQDQQRQLVDYISKRSKQFKPVKLKHETKFVAFSTCENFSTDNRVIVVGTIQ 590
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2390

A DNA sequence (GASx129) was identified in *S.pyogenes* <SEQ ID 7279> which encodes the amino acid sequence <SEQ ID 7280>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -6.05   Transmembrane    5 -  21 (   4 -  22)
   INTEGRAL    Likelihood = -5.04   Transmembrane  191 - 207 ( 186 - 209)

----- Final Results -----
              bacterial membrane --- Certainty=0.3421(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>

LPXTG motif: 181-186
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC97151 GB:U49397 unknown [Streptococcus pyogenes]
 Identities = 64/213 (30%), Positives = 106/213 (49%), Gaps = 20/213 (9%)

Query: 1   MKKSILRILAIGYLLMSFCLLDSVEAENLTASINHWINQVDVATNKQSSDIDETFMFVI 60
           M+K    + ++   +L        +V A++ T    +I V N ++ A +          F    +
Sbjct: 1   MRKYWKMLFSVVMMLTMLAPNQTVLAKDSTVQTSISVRNVLERAGDSTF------FSIAL 54

Query: 61  EALDKESPLPNSVTTSVKGNGKTSPEQLTPSEVGQYHYKIIQLLGKNSQYHYDETVYEVV 120
           E++D    +       ++ G+GK SP  L F+ VGQY Y+++Q    +N  Y  D TV++V+
Sbjct: 55  ESIDAMKTIEE---ITIAGSGKASPSPLNFTTVGQYTYRVYQKPSQNKUYQADTTVPDVL 111

Query: 121 IYVLYNEQSGALETNLVSNKLGSTEKSELIFKQEYSEKTPEPHQPDTTEKEKPQKKRNGI 180
           +YV Y+E  G L    ++S + G+ EKS + FK +    K    P QPD  +
Sbjct: 112 VYVTYDR-DGTLVAKVISRRAGDEEKSATTFKPKRLVKPIPPRQPDIPKTP--------- 161

Query: 181 LPSTGEMVSYVSALGIVLVATTTLYSIYKKLKT 213
           LP  GE+ S +  L IVL+  + L  + KKLK+
Sbjct: 162 LPLAGEVKSLLGILSIVLLGLLVLLYV-KKLKS 193
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2391

A DNA sequence (GASx130R) was identified in *S.pyogenes* <SEQ ID 7281> which encodes the amino acid sequence <SEQ ID 7282>. Analysis of this protein sequence reveals the following:

```
Possible site: 57

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.1614(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB54046 GB:AJ245436 hypothetical protein, 57.8 kD [Pseudomonas
           putida]
 Identities = 128/388 (32%), Positives = 204/388 (51%), Gaps = 21/388 (5%)

Query: 4   IGSVVQRQELVFIPAQLKRINEVQHAYKCQTCSDNSLSDKIIKAPVPKAPLAHSLGASSI 63
           IG  V  Q L  +P Q++ I HV+  Y C+ C     ++      A  P   +  S+ S S+
Sbjct: 126 IGREVSEQ-LEIVPMQIRVIKEVRKVYGCRDCESAPVT-----ADKPAQMIEKSMASPSV 179

Query: 64  IAHTVHQKPTLKVPNYRQEEDWNKLGLSISRKEIANWHIKSSQYYPEPLYDLLRDILLSQ 123
           +A  +  K+    +P +R E+   + G+ I R+ +A W I+ S++ F+PL +L+R+ LL+
Sbjct: 180 LAMLLTTKYVDGLPLHRFEKVLGRHGIDIPRQTLARWVIQCSEH-PQPLLNLMRESLLNS 238

Query: 124 EVIHADETSYRVLESD----TQLYTYWTFLSGKHEKKGITLYHHDKRRSGLVTQEVLGDY 179
            +IH DET  +VL+     +  W     G  ++   + L+ +   R+  V   +L  Y
Sbjct: 239 RIIHCDETRVQVLKEPGREPSSQSWMWVQTGSPPDRP-VILFDYATSRAGEVPVRLLDGY 297

Query: 180 SGYVHCDMHGAYRQL---EHAKLVGCWAHVRRKFFEATPKQAD-KTSLGRKGLVYCDKLF 235
            GYV D +  Y  L   +  +GCWAH RRKF RA   Q  KT     L  +KL+
Sbjct: 298 RGYVMTDDYAGYNALAAQDGLERLGCWAHARRKFVEAQKVQPKGKTGRADIALNLINKLY 357

Query: 236 ALEAEWCELPPQERLVKRKEILTPLMTFFDWCR--EQVVLSGSKLGLAIAYSLKHERTF 293
           +E +  +  +R V R E   PL+T   +W     + V + + LG AI Y   +
```

```
Sbjct: 358 GVERDLKDSDDEDRKVARMERSLPLLTQLKNWVEKTQPQVTTQNALGKAIGYLASNWSKL 417

Query: 294 RTVLEDGHIVLSNNMAERAIKSLVMGRKNWLFSQSPEGAKAAAIIMSLLETAKRHGLNSE 353
           +E G++ + NN AERAI+  V+GRKNWLFS + +GA A+A + SL+ETAK +G
Sbjct: 418 ERYVEHGYLPMDNNAAERAIRPFVIGRKNWLFSDTPKGATASAQLYSLVETAKANGQEPY 477

Query: 354 KYISYLLDRLPNEETLAKREVLEAYLPW 381
           ++ + L+RLP   ++   E  EA LPW
Sbjct: 478 AWLRHALERLPQACSV---EDYEALLPW 502
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2392

A DNA sequence (GASx131R) was identified in *S.pyogenes* <SEQ ID 7283> which encodes the amino acid sequence <SEQ ID 7284>. Analysis of this protein sequence reveals the following:

```
Possible site: 37

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4465(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2393

A DNA sequence (GASx132R) was identified in *S.pyogenes* <SEQ ID 7285> which encodes the amino acid sequence <SEQ ID 7286>. Analysis of this protein sequence reveals the following:

```
Possible site: 46

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1529(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA84885 GB:AB024946 orf50 [Escherichia coli]
 Identities = 37/91 (40%), Positives = 53/91 (57%)

Query: 10  QVYLVCGKTDMRQGIDSLAYLVKSQHELDLFSGAVYLPCGRRDRFKALYWDGQGFWLLY 69
           ++++LV G TDMR G + LA  V++  + D FSG +++F G R D+ K L+ D  G  L
Sbjct: 9   RIWIVAGITDMRNGFNGLASKVQNVLKDDPFSGHGFIFNGRRGDQIKVLWADSDGLCLFT 68

Query: 70  KRPENGKLAWPRNRDEVECLTAVQVDWLMKG 100
           KR E G+  WP  RD     L+T  Q+  L++G
Sbjct: 69  KRLERGRFVWPVTRDGKVHLTPAQLSMLLEG 99
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2394

A DNA sequence (GASx133R) was identified in *S.pyogenes* <SEQ ID 7287> which encodes the amino acid sequence <SEQ ID 7288>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1979(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2395

A DNA sequence (GASx135R) was identified in *S.pyogenes* <SEQ ID 7289> which encodes the amino acid sequence <SEQ ID 7290>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
              bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2396

A DNA sequence (GASx136) was identified in *S.pyogenes* <SEQ ID 7291> which encodes the amino acid sequence <SEQ ID 7292>. Analysis of this protein sequence reveals the following:

```
Possible site: 54

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-11.73   Transmembrane  222 - 238 ( 212 - 242)
   INTEGRAL    Likelihood =-10.88   Transmembrane   37 -  53 (  32 -  57)
   INTEGRAL    Likelihood = -9.87   Transmembrane  462 - 478 ( 456 - 478)
   INTEGRAL    Likelihood = -4.25   Transmembrane  119 - 135 ( 117 - 137)
   INTEGRAL    Likelihood = -2.60   Transmembrane  308 - 324 ( 306 - 324)
   INTEGRAL    Likelihood = -1.28   Transmembrane  164 - 180 ( 164 - 180)
   INTEGRAL    Likelihood = -0.06   Transmembrane  137 - 153 ( 137 - 153)
   INTEGRAL    Likelihood = -0.06   Transmembrane  343 - 359 ( 343 - 359)
```

```
----- Final Results -----
                bacterial membrane --- Certainty=0.5692(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04077 GB:AP001508 short-chain fatty acids transporter
            [Bacillus halodurans]
 Identities = 158/465 (33%), Positives = 248/465 (52%), Gaps = 41/465 (8%)

Query: 15  IKTEKRPMDRYIDGPMKWMPESLPICPILTPLVVTMSVLMTDSPPIGTEKTGGIIYGWVN 74
           I    R M RY+       P+      +LTPLV  +S++ T+S    T  T  I+  W
Sbjct: 5   ISLSNRLMQRYL-------PDPFLPVVLLTPLVPALSLIFTES----TPLT--IVQYWGE 51

Query: 75  GFWGLLSFAMQMTILLATGNAVASSPPAHKMFKSLAKLPQTRTQIPIFSIVVGSIPGFLH 134
           GFWGLLSF+MQM ++L TG+ +ASSP   K   +LA LP +  Q +   VV  +  F++
Sbjct: 52  GFWGLLSFSMQMVLVLVTGHVLASSPLFKKGLGALAGLPASPGQAILLVTVVSLVASFIN 111

Query: 135 WGLGMMVAIVPGKELLVQARQKGIKVHTPLFVATLFPTFLPATSGLSGAAVLYSATPDYL 194
           WG G+++  +F KEL     +K   V   L +A+ +  F+    GLSG+  L  ATPD+
Sbjct: 112 WGPGLVIGALFAKELA----KKVDNVDYRLLIASAYSGPFMIWHGGLSGSVPLTIATPDHF 167

Query: 195 RNSVADAYKQVVPESVPLTESVL---NLPPISLLVVCMLVPLCFALLAHPKDETKIME-- 249
                +          +P +E++    NL  +  L+    +PL   L+   K +T  ++
Sbjct: 168 AQDMIGV--------IPTSETIFAFYNLAIVFALFIA--IPLANRLMMPGKSDTVTVDRS 217

Query: 250 -LDDEIYHHSLDTASHVVIARNTPAEKMNASRLVMYLVGGAIVSYSLYHFSVVGLSGLDL 308
            LDD      L  AS + +   TP++++  SR++  LVG    +  Y+F+  G   L+L
Sbjct: 218 LLDDG---RDLQAAS-LELEAMTPSDRLENSRMISLLVGVLGLVFLGYYFATNGPE-LNL 272

Query: 309 NCFNFLFLGLGLLLCGQQGPEYYGSLFKDGVMSSWGLVLQFPFYYAGIFGIIQSTGLGLEI 368
           +  N LFL LG+L  G   P+ +       V  + G+++QPPFYAG+ GI+ S+GL   +
Sbjct: 273 DIVNSLFLFLGILPHGT--PKLFLKAVTSAVKGASGIIIQFPFYAGLMGIMVSSGLATVM 330

Query: 369 SHFPVAISNGTTWDVFAYLYSALLNIAVPSQGSKFVIEAPYIVPATIEVGNDLGKILQAY 428
           S  PV+ SN  T+P+F +L + ++N+ VPSQG ++ ++AP ++ A   +G    K   A
Sbjct: 331 SEAPVSFSNEVTFPLFVFLSAGIVNVFVPSQGGQWAVQAPVVLEAAQSLGVPAAKAAMAV 390

Query: 429 QLGDATTNLIVPFWALSYLSNFKLKPNQIVAYTIPCVLVVTGIAI 473
             GDA TN+I PFWAL  L+   LK   I+ + +  +LVV+G+ I
Sbjct: 391 AWGDAWTNMIQPFWALPALAIAGLKAKDIMGPCV-MILVVSGVVI 434
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2397

A DNA sequence (GASx137R) was identified in *S.pyogenes* <SEQ ID 7293> which encodes the amino acid sequence <SEQ ID 7294>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.2591(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC22434 GB:U32761 transcriptional regulator [Haemophilus influenzae Rd]
 Identities = 37/107 (34%), Positives = 56/107 (51%), Gaps = 1/107 (0%)

Query: 21  LHRQNLVTFDKTFMINHQLTTLFEEANSLPVVKCYSASWDFLLNCTRYS-SYLTTLLPRPI 79
            LH+Q +  FD+TFMI+H L    FE  N  P +   S+ WDFLL+  + +   LYILP P+
Sbjct: 205 LHQQKMAIFDQTFMIHHHLKRAPERNNCYPDIVLDSSCWDFLLSAVKTNKELLTILPLPM 264

Query: 80  THFAHMDGLVEVQLTKHPKWEVVLASLKHNKTSHELKHYIKHTILDYF 126
             H    +  ++     W+V L   +    +HL+ YI   +L+ F
Sbjct: 265 AELYHSKEPLCRKIESPVPWKVTLCRQRKTVYTHLEEYIFDKLLEAF 311
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2398

A DNA sequence (GASx140) was identified in *S.pyogenes* <SEQ ID 7295> which encodes the amino acid sequence <SEQ ID 7296>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3351(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
!GB:U32761 acetate CoA-transferase, alpha subunit [H...    215  4e-55
 Identities = 105/213 (49%), Positives = 146/213 (68%)

Query: 22  ENKRIAIAEAISHIKDGDTIMVGGFMANGTPEALIDALVDKGTKDLTLICNDAGFVDRGV 81
           + K + + +A    +DG TIMVGGFM  GTP  L++AL++ G +DLTLI ND  FVD G+
Sbjct: 2   KTKLMTLQDATGFFRDGMTIMVGGFMGIGTPSRLVEALLESGVRDLTLIANDTAFVDTGI 61

Query: 82  GKMVANHQPKTIYATHIGLNKEAGRQMTAGETTIELIPQGTFAEKIRIGAYGIGGFYTPT 141
           G ++ N + + + A+HIG N E GR+M +GE  + L+PQGT  E+IR G  G+GGF TPT
Sbjct: 62  GPLIVNGRVRKVIASHIGTNPETGRRMISGEMDVVLVPQGTLIEQIRCGGAGLGGFLTPT 121

Query: 142 GVGTLVAEGKETKTIKGKTYLLEYPFEADVALIFAMQADEMGNLQYSGSENNFNQLMAAC 201
           GVGT+V EGK+T T+ GKT+LLE P  AD+ALI A++ D +GNL Y  S  NFN L+A
Sbjct: 122 GVGTVVEEGKQTLTLDGKTWLLERPLRADLALIRAHRCDTLGNLTYQLSARNFNPLIALA 181

Query: 202 AKTTIVQARSIVPVGTIQPECVHTPHIFVDYIV 234
           A  T+V+  E+V  G +QP+ + TP    +D+I+
Sbjct: 182 ADITLVEPDELVETGELQPDHIVTPGAVIDHII 214
           subunit (EC 2.8.3.-). [Escherichia coli]
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2399

A DNA sequence (GASx141) was identified in *S.pyogenes* <SEQ ID 7297> which encodes the amino acid sequence <SEQ ID 7298>. Analysis of this protein sequence reveals the following:

```
Possible site: 41

>>> Seems to have no N-terminal signal sequence
```

```
----- Final Results -----
                bacterial cytoplasm --- Certainty=0.4941(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF12248 GB:AE001862 CoA transferase, subunit B [Deinococcus radiodurans]
 Identities = 114/203 (56%), Positives = 158/203 (77%), Gaps = 3/203 (1%)

Query: 11  QNRIAKRVAKRLEDGTLVNLGIGLPTKVANFVPKEMTVYFQSENGFIGLGP--KSDDPNS 68
           ++ +A R A+EL+DG  VNLGIGLPT VAN +P  M+V+ QSENG +G+GP    D+ +
Sbjct: 5   RDEMAARAAQELQDGYYVNLGIGLPTLVANHIPAGMSVWLQSENGLLGIGPFPTEDEVDP 64

Query: 69  TIVNAGSQPVTVYPGAAFFNSADSFGIIRGGHVDLTVLGALEIAENGDIANYLIPGKMVP 128
            ++NAG Q VT  PGA+FF+SADSF +IRGGHV+L +LGA++++E GD+AN++IPGKMV
Sbjct: 65  DLINAGKQTVTALPGASFFSSADSFAMIRGGHVNLAILGAMQVSETGDLANWMIPGKMVK 124

Query: 129 GMGGAMDLLVGAKKVIVAMEHTNKG-KHKLLKECTLPLTAKGVVDLIITEMGVFKVTPDG 187
           GMGGAMDL+ G ++V+V MEH  KG  HK+L+ECTLPLT +GVVD IIT++GV  VTP G
Sbjct: 125 GMGGAMDLVAGVQRVVVLMEHVAKGDAHKILRECTLPLTGQGVVDRIITDLGVLDVTPCG 184

Query: 188 IQVIEISEGFTFDEVQAATGVPL 210
           ++++E++ G T DE++  TG  +
Sbjct: 185 LKLVELAPGVTLDELRQKTGADI 207
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2400

A DNA sequence (GASx144) was identified in *S.pyogenes* <SEQ ID 7299> which encodes the amino acid sequence <SEQ ID 7300>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.3227(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA29948 GB:AP000003 137aa long hypothetical protein [Pyrococcus
          horikoshii]
 Identities = 49/113 (43%), Positives = 71/113 (62%), Gaps = 1/113 (0%)

Query: 5   PEPWGPYSTYTIEGHFLYTAGQLPLNEVTQQLSDG-PEAQCRQVFVNLQSILAEQKLDLN 63
           P+P+GPYS     G+FL+ AGQ+P++P TG++  G  + Q RQV  N+++IL      LN
Sbjct: 22  PKPIGPYSQAIKAGNFLFIAGQIPIDPKTGEIVEGDIKDQTRQVLENIKAILEAAGYSLN 81

Query: 64  HIYKLNVYLTDVTNVEILNHVWTDLFSEPYPVRTAVQVSALPLQALIEVEAVA 116
            + K+ VYL D+ +   +N V  + F E  P R AV+VS LP   LIE+EA+A
Sbjct: 82  DVIKVTVYLKDWNDFAKMNEVYAEYFGESKPARVAVEVSRLPKDVLIEIEAIA 134
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2401

A DNA sequence (GASx146) was identified in *S.pyogenes* <SEQ ID 7301> which encodes the amino acid sequence <SEQ ID 7302>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1238(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2402

A DNA sequence (GASx147) was identified in *S.pyogenes* <SEQ ID 7303> which encodes the amino acid sequence <SEQ ID 7304>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-11.46   Transmembrane  456 - 472 ( 452 - 481)
   INTEGRAL    Likelihood = -8.17   Transmembrane  603 - 619 ( 595 - 623)
   INTEGRAL    Likelihood = -6.85   Transmembrane  495 - 511 ( 491 - 518)
   INTEGRAL    Likelihood = -5.31   Transmembrane  420 - 436 ( 418 - 443)
   INTEGRAL    Likelihood = -4.99   Transmembrane  396 - 412 ( 392 - 413)
   INTEGRAL    Likelihood = -1.59   Transmembrane  522 - 538 ( 522 - 538)
   INTEGRAL    Likelihood = -0.64   Transmembrane  577 - 593 ( 577 - 593)
   INTEGRAL    Likelihood = -0.43   Transmembrane  377 - 393 ( 377 - 393)

----- Final Results -----
              bacterial membrane --- Certainty=0.5585(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA04270 GB:D17462 Na+ -ATPase subunit I [Enterococcus hirae]
 Identities = 232/681 (34%), Positives = 370/681 (54%), Gaps = 40/681 (5%)

Query: 1   MAISCMKKLAMVFEKDYLDLVLKTLQQSGLVEVRDMKQLKH---WDDAFNKGNVKLPQIV 57
           MA+++M+K+ ++ +K   +++L+ +Q    VE+RD+ Q      W + F        P+++
Sbjct: 1   MAVTKMEKVTLISDKKNREILLQAVQGLHAVEIRDLFQESENNQWVETF----FPEPEMI 56

Query: 58  QYDLTHQKPLLDDEALQYLLQSQQELSNSLASLSAFLPPIGKLTALRQ--KTPSLSFKQF 115
           D   K      L Y L          + +   F+  G+ +  +Q  K   LS
Sbjct: 57  DKDKELAK-------LSYKLTD-------IKTAIQFIEHGEKSQKKQHLKRRELSLDTL 102

Query: 116 EERHRQQAAQTALKWMSQKIERLEQLQSKIDQLTEYCQELEKWRSLTVLPQDLAQFHFLS 175
           E+ + ++A   L+ +    E+ EQL  + QL +    L  W++L + P+
Sbjct: 103 EKNYSEEAFSKKLSEVLLLKEQWSQLVDERQQLEDQENWLLNWQNLDLAPKAFDS-QMTK 161

Query: 176 ARVGTIPSTANNHFYHQLKQHKGLPIEEVYH----TEFEYGLVLFWQAQDTTHLQKYQFK 231
           +GT+ +    F  ++ +    ++EE+    T F Y ++   +++      +Y F
Sbjct: 162 LVIGTVNAKNAESFKAEVAEINEAYLEEINSSPTTYFAYIVLRADESRMEETASKYGFV 221
```

```
Query: 232 PLLYKSQLLPSEQLRINKELLTNWLAKDSLLKRLRQSQKILAQLCVEIDYVLSQVQKQQ 291
             Y  +  P +QL   K+ L      ++  L  +       +   +   +  L++ +R+
Sbjct: 222 KEDYLYEGTPQCQLVAAKQSLQEIKDQQKKLSSAIGACSGYIKDFEWTEEIFLARSERBA 281

Query: 292 TKKQLLGTRHLIALEGWIEADSVNQLKGLMTKTLGDMFYLDSYDVTPDDW--EDVPIKLR 349
             K +++ T +LI ++GN++ +   +L  ++   L       ++D   D+   E+VP KL+
Sbjct: 282 IKDRIIHTPYLLILQGNVDHERKQELIMALQNILASEKVYLTPDEPTDNEIASEVPTKLK 341

Query: 350 NHRYIAPFELVTEMYALPKYQEKDPTPPLAPLYLTFPGMMVADLGYGLLLVAVTLAALVF 409
             NH  +APFR++TEMY+LPKY+E DPTP++ P YL FPGMMVADFGYGLL++
Sbjct: 342 NHPIVAPFEMLTEMYSLPKYEEVDPTPWMMPFYLVFPGMMVADIGYGLLMPLGAFLLQKL 401

Query: 410 FNLQKTSKRLVTPFNILAISVAIWGLIYGSPFG---------FDLPVALLSTKTDVTTIL 460
             L +  +R    FF ILAI    IWG IY SPFG           LP  +LST  DV TIL
Sbjct: 402 VVLPRGMQRFAKFPEILAIPSIIWGPIYGSFPGAALPKBIPGIHLPFPILSTTDDVNTIL 461

Query: 461 VVSLLFGFVTLIFGLLLGAWQQVRMKAYATVTSSLAWTFILLGLLLFILGKNVSGLAYL 520
             ++S++PG + ++ GL + A + ++ KAY  A      AW +ILLG++L +LG
Sbjct: 462 ILSVIFGLIQILVGLFIAAKESHIKKKAIVDAVNDGPAWQWILLGIILILLLGTMTLKNNAF 521

Query: 521 SVIGKWLALGNAFGILVVSLLKSKSLL-GLGSGLYNLYGISSYLSDLVSFTRLMALGLSG 579
              +G  LA+ +A  IL++ + +S S   G+  G YNLYG++ Y+ DLVS+TRLMALG+SG
Sbjct: 522 VYLGGALAVLSAVCILIIPVPQSSSKAKGIAKGAYNLYGLTGYIGDLVSTTRLMALGISG 581

Query: 580 ASIGAAFNMIVGIFPPVTRFTVGIFIFILLHAINIFLSMLSGYVHGARLIFVEFPGKFYE 639
             SI AAFNM+V   PP  RF+VGI + I+L A+N+FL++LS YVHGARL +VEFPGKFY
Sbjct: 582 GSIAAAFNMLVAFMPPAARFSVGILLIIVLQAINMFLTLLSAYVHGARLQYVEFPGKFYT 641

Query: 640 GGGKAFNVLKLADNYVNVNES 660
             GGG++F FLK  + YVN+N +
Sbjct: 642 GGGRSFKPLKTVEKTVNINHK 662
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2403

A DNA sequence (GASx148) was identified in *S.pyogenes* <SEQ ID 7305> which encodes the amino acid sequence <SEQ ID 7306>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -7.80   Transmembrane   28 -  44 (  21 -  51)
   INTEGRAL    Likelihood = -6.85   Transmembrane  148 - 164 ( 146 - 170)
   INTEGRAL    Likelihood = -2.81   Transmembrane  105 - 121 ( 105 - 123)

----- Final Results -----
               bacterial membrane --- Certainty=0.4121(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA03841 GB:D16334 Na+-ATPase K subunit [Enterococcus hirae]
 Identities = 85/150 (56%), Positives = 107/150 (70%)

Query: 20  HYFTAHGGVFFAALGIVLAVALSGMGSAYGVGKGQQAAAALLKEEPEKFTSALILQLLPG 79
            +  T +GG+ FA L +  A   SG+GSA GVG  G+AAAAL   +PEKF  ALILQLLPG
Sbjct: 4   YLITQNGGMVFAVLAMATATIFSGIGSAKGVGMTGRAAAALTTSQPEKFQAALILQLLPG 63

Query: 80  SQGIYGFAIGILIWMKLTPELSVNQGLAYFLVSLPIAIVGYFSAKHQGNVSVAGMQILAK 139
            +QG+YGF I  LI++ L  ++SV QGL +   SLPIA  G FS   QG V+ AG+QILAK
```

```
Sbjct: 64  TQGLYGFVIAFLIPINLGSDMSVVQGLNFLGASLPIAPTGLFSGIAQGKVAAAGIQILAK 123

Query: 140 RPKDFMKGVILAAMVETYAILAFVVSFILL 169
           +P+   KG+I AAMVETYAIL FV+SF+L+
Sbjct: 124 KPEHATKGIIFAAMVETYAILGFVISFLLV 153
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2404

A DNA sequence (GASx149) was identified in *S.pyogenes* <SEQ ID 7307> which encodes the amino acid sequence <SEQ ID 7308>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4510(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA04272 GB:D17462 Na+ -ATPase subunit E [Enterococcus hirae]
 Identities = 43/193 (22%), Positives = 95/193 (48%), Gaps = 2/193 (1%)

Query: 1   VNDITQLRQNVLEKAHQEGQQCLKIATDSLDTDPKRRQQQSLHDLKAKRQKSLKALEQQF 60
           V+ I ++   + E A  E      ++     +D  F+ ++ Q   D + ++  +L+ +E+ +
Sbjct: 3   VDAIDKIITQINETAQLERASFEEMKRKEIDKFEVKKNQIEADPQKRKASKLEEIERSY 62

Query: 61  QVAQQQLKNQERQALLALKQDSIKELFEASLEKMTNFSKEERLAFLKQVLSKYP-EQPLQ 119
           +  + + K Q +Q +L  KQ+ ++ LF  +  ++ N  KEE+LA +KQ++    P     +
Sbjct: 63  RQLRNKQKMQVKQEILNAKQEVLQRLFTEATLQLENEPKEEQLALMKQMIQTLPINTAR 122

Query: 120 VTFGEKTGQKPSSYDCAELRLAPPQLSYNQELIPQ-EAGFLVSLDQVDDNYLYRYLLESV 178
           +  GEK+    +   AE     P      ++   + +AG ++    + N+L+ +L++ +
Sbjct: 123 LIPGEKSADILTPAVIAEWNEELPFELIREDFTEKAQAGLIIDDAGIQYNFLPSHLIKEI 182

Query: 179 LKEESSRIIDMLF 191
            + S+ I   LF
Sbjct: 183 QETMSARIAKELF 195
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2405

A DNA sequence (GASx150) was identified in *S.pyogenes* <SEQ ID 7309> which encodes the amino acid sequence <SEQ ID 7310>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3095(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA04273 GB:D17462 Na+ -ATPase subunit C [Enterococcus hirae]
 Identities = 94/326 (28%), Positives = 167/326 (50%), Gaps = 5/326 (1%)

Query: 6   ELNTTISVKEKELLTYEQFDKLLQAPNTTTLARLLHQSVTHLTVDDLNDLDRLESILMAE 65
            ELN  I  +E EL++K+ F++++Q  +  +L  +L  ++Y   + D  D D  E+ L  E
Sbjct: 5   KLNPLIRGKELELISKDTFKQMIQTDSIDSLGKILQSTIYQPYIYDGFDKD-PEAMLSQE 63

Query: 66  LTKTYRWAFAETPQPDIVQLPTLRYTYHNVKVLLKAKASQADLSHLLLPIGDKPLVALEH 125
             +K ++W      P+P+IV ++T+RYT+HN+KVL KA+ +  +L HL +  G   L  L+
Sbjct: 64  RSKLFQWLKESAPEPEIVWIYTMRYTFHNLKVLTKAEITGQNLDHLYIHDGFYSLEVLKD 123

Query: 126 LIRTWTSDEFPKEVVTEIQSIWAEYQDYQDIRVLEIGTDLAYFKALKQIAQRLEDPVPQQ 185
             I T  S E P  ++  I+ +     ++   ++ +++  D  +     +++ ++L  P  +
Sbjct: 124 AIHTQVSVELPDSLMDYIREVHEYCESSTILQGIDVIYDRCPLTEQRRLGEQLGYPELLE 183

Query: 186 AVLIVIDLVNLLTVRRAKSQNKPISFWMQLLSDEASRPSKTFITLEDDKDLMTWPENVTP 245
             ++  IDL N+ T  R   Q++     FM  ++S  S P  T  ++      ++ ++ + +
Sbjct: 184 EIIAPIDLTINITTARGILQERSAGFMTTVISSSGSIPKDTLLSFVRG-EWASFTQFLLT 242

Query: 246 DSYMTALKPYSEKLRQGTLQTTELEYLVDECLYHLFAKAKYQVDGPYVLARFLLAKSFEV 305
             Y   LK   + +  +     LE L D+ L   +  A+ Q  GP  L  FL AK  E
Sbjct: 243 TDYSELLK---QVIHEEQIDLVSLEQLKDDYLSSFYQVAQTQAPGPLPLLAFLNAKEVES 299

Query: 306 KNLRLLAAALANDLPKERVIEEMMRPI 331
             KNLRLL     N    E++ ERMR +
Sbjct: 300 KNLRLLIIGKRNHFSLEQLKKRMRQV 325
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2406

A DNA sequence (GASx151) was identified in *S.pyogenes* <SEQ ID 7311> which encodes the amino acid sequence <SEQ ID 7312>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0484(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA04274 GB:D17462 Na+ -ATPase subunit G [Enterococcus hirae]
 Identities = 45/101 (44%), Positives = 65/101 (63%)

Query: 6   YKVGVIGNRDVILPPQMIGPQTPFVIKPQDAINQLRQLAMEDPGIIYITEDIAAAIPEAL 65
            YK+GV+G+D + PF++ GF        +   ++A  ++G+IYITE  A  +PE +
Sbjct: 3   YKIGVVGDKDSVSPPRLRGPDVQHGTTKTKIRKTIDKMAKNEYGVIYITEQCANLVPESTI 62

Query: 66  THYDNQVLPAVIPLPTHQGAQGIGLSRIQAMVEKAVGQNIL 106
             Y  Q+ PA+I +P+HQG  GIGL  IQ  VEKAVGQNIL
Sbjct: 63  ERYKGQLTPAIILIPSHQGTLGIGLEEIQNSVEKAVGQNIL 103
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2407

A DNA sequence (GASx152R) was identified in *S.pyogenes* <SEQ ID 7313> which encodes the amino acid sequence <SEQ ID 7314>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1048(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2408

A DNA sequence (GASx156) was identified in *S.pyogenes* <SEQ ID 7315> which encodes the amino acid sequence <SEQ ID 7316>:

```
EYSIIPQLKETIHYISLKLEEAERASLVRIMKITS
```

Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.5026(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA04277 GB:D17462 Na+ -ATPase subunit D [Enterococcus hirae]
 Identities = 119/201 (59%), Positives = 151/201 (74%), Gaps = 2/201 (0%)

Query: 10  RLNVKPTRMELSNLKNRLKTTATRGHKLLKDKRDELMRRFVDLIRENNELRQTIEKELAAN 69
           RLNV PTRMEL+ LK +L TATRGHKLLKDK+DELMR+F+ LIR+NNELRQ IEKE
Sbjct: 2   RLNVNPTRMELTRLKKQLTTATRGHKLLKDKQDELMRQFILLIRKNNELRQAIEKETQTA 61

Query: 70  MKEPVLAKASENSLMVEELFAVDVHEVTLWDIIENIMSVNVPKFHVQSNTAREQEQGEFA 129
           MK+PVLAK++       ++EL A+P   V++ +  +NIMSV VP  + Q +     +   E
Sbjct: 62  MKDPVLAKSTVEEAFIDELLALPAENVSISVVEKNIMSVKVPLMNFQYDETLNETPLE-- 119

Query: 130 YGYLGSNSENDNTIQKTKELLEKLLRLAEVEKTCQLMADDIEKTRRRVNGLEYSIIPQLK 189
           Y YL SN+E+D +I     +LL KLL+LAEVEKTCQLMA++IEKTRRKVN LEY  IPQL+
Sbjct: 120 YGYLHSNAELDRSIDGFTQLLPKLLKLAEVEKTCQLMAKEIEKTRRRVNALEYMTIPQLR 179

Query: 190 ETIHYIELKLEEAERASLVRI 210
           ETI+YI++KLEE ERA + R+
```

```
Sbjct: 180 ETFYYIKMKLEENERAEVTRL 200
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2409

A DNA sequence (GASx161R) was identified in *S.pyogenes* <SEQ ID 7317> which encodes the amino acid sequence <SEQ ID 7318>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2410

A DNA sequence (GASx164) was identified in *S.pyogenes* <SEQ ID 7319> which encodes the amino acid sequence <SEQ ID 7320>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.06    Transmembrane     9 -  25 (   9 -  25)

----- Final Results -----
               bacterial membrane --- Certainty=0.1426(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

A related sequence was also identified <SEQ ID 9091> which encodes the amino acid sequence <SEQ ID 9092>. Analysis of this protein sequence reveals the following:

```
      Possible cleavage site: 33
>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty= 0.300(Affirmative) < succ>
               bacterial membrane --- Certainty= 0.000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty= 0.000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2411

A DNA sequence (GASx165) was identified in *S.pyogenes* <SEQ ID 7321> which encodes the amino acid sequence <SEQ ID 7322>. Analysis of this protein sequence reveals the following:

```
Possible site: 59

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2251(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2412

A DNA sequence (GASx166) was identified in *S.pyogenes* <SEQ ID 7323> which encodes the amino acid sequence <SEQ ID 7324>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
              bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2413

A DNA sequence (GASx167) was identified in *S.pyogenes* <SEQ ID 7325> which encodes the amino acid sequence <SEQ ID 7326>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
              bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2414

A DNA sequence (GASx168R) was identified in *S.pyogenes* <SEQ ID 7327> which encodes the amino acid sequence <SEQ ID 7328>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2415

A DNA sequence (GASx169R) was identified in *S.pyogenes* <SEQ ID 7329> which encodes the amino acid sequence <SEQ ID 7330>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2416

A DNA sequence (GASx170) was identified in *S.pyogenes* <SEQ ID 7331> which encodes the amino acid sequence <SEQ ID 7332>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -2.34   Transmembrane  154 - 170 ( 153 - 170)
   INTEGRAL    Likelihood = -1.12   Transmembrane   20 -  36 (  19 -  36)
   INTEGRAL    Likelihood = -0.69   Transmembrane   52 -  68 (  52 -  68)
   INTEGRAL    Likelihood = -0.53   Transmembrane  399 - 415 ( 399 - 415)

----- Final Results -----
              bacterial membrane --- Certainty=0.1935(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB05347 GB:AP001512 cystathionine beta-lyase [Bacillus halodurans]
 Identities = 200/384 (52%), Positives = 262/384 (68%), Gaps = 3/384 (0%)

Query: 79  IAEVYEMRENTTLLHGYTVIDEFTGAASVPIYQTSTFHNSELYCPSQKHLVTRFSNPTTR 138
           ++E Y ++  T LLH     +D+ TGA SVPI    STFH  + +      + Y+R  NPT +
Sbjct: 1   MSEQYSLQ--TKLLHNEHKVDQATGAVSVPIQHASTFHQFD-FDTFGTYDYSRSGNPTRD 57

Query: 139 ALEDGLACLEKATYAVAYASGMAAISTVLMLLKAGDHVIFPLEVYGGTCQFATAILPNYQ 198
           ALE  +A LE   + A+ASGMAAIST  MLL  GDHV+    +VYGGT +  T +L
Sbjct: 58  ALEAAIAELRGGNHGFAFASGMAAISTAFMLLSKGDHVVLTXDVYGGTFRLVTEVLTRIG 117

Query: 199 IETSFVDMADLATVKASIRPNTRMIYLETPSNPLLKICDISELVQLAKAYGVLTVADNTF 258
           IE +FVEM +LA V A+IRPNTR++Y+ETPSNP L I DI  +V LAK +  LT  DNTF
Sbjct: 118 IEHTFVEMTNLAEVAAAIRPNTRVLYMETPSNPTLNITDIRGVVSLAKEHECLTFLDNTF 177

Query: 259 MTSLYQEPLAMGVDIVVESVTKFINGHSDVVAGLAATNNEAIYNQLKLFQKNFGAIVGVE 318
           +T   Q PL +GVD+V+ S TKFI GHSDVVAGLA T NE +  +L   Q +FGAI+GV+
Sbjct: 178 LTPALQRPLELGVDVVLHSATKFIGGHSDVVAGLATKNEELGKKLAFLQNSFGAILGVQ 237

Query: 319 DAWLILRGMKTMGIRMEQAVKNAQQLANYLAKHPKVLKVHYPGLDSHPNHDTHLQQAKNG 378
           D WL+LRG+KT+ +RME   K AQQ+A +L   P+V +V+YPGL  HP H+   +QA+
Sbjct: 238 DVWLVLRGLKTLHVRMEHGEKGAQQIAEWLQSVPEVKRVYYPGLKDHPGHELQKRQAEGF 297

Query: 379 GAVLSFELASKEELMTFTHRIQLPILAVSLGGVESILSHPATMSHACLSPQARLEQGVVD 438
           GAVLSFEL ++E +  F   ++LP+ AVSLG VESILS+PA MSHA +  + R  +G+ D
Sbjct: 298 GAVLSFELENEEAVRRFVEHVKLPVFAVSLGAVESILSYPAKMSHAAMPKEEREARGIRD 357

Query: 439 GLLRLSCGVENIEDLLADFEQALA 462
           GLLRLS G+E  E+L+ADF+ A A
Sbjct: 358 GLLRLSVGLEKPEELMADFKAAFA 381
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2417

A DNA sequence (GASx178) was identified in *S.pyogenes* <SEQ ID 7333> which encodes the amino acid sequence <SEQ ID 7334>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1492(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2418

A DNA sequence (GASx182) was identified in *S.pyogenes* <SEQ ID 7335> which encodes the amino acid sequence <SEQ ID 7336>. Analysis of this protein sequence reveals the following:

```
Possible site: 22
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2584(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2419

A DNA sequence (GASx187) was identified in *S.pyogenes* <SEQ ID 7337> which encodes the amino acid sequence <SEQ ID 7338>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2084(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2420

A DNA sequence (GASx188) was identified in *S.pyogenes* <SEQ ID 7339> which encodes the amino acid sequence <SEQ ID 7340>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2060(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAG05515 GB:AE004640 conserved hypothetical protein [Pseudomonas aeruginosa]
 Identities = 140/442 (31%), Positives = 208/442 (46%), Gaps = 73/442 (16%)

Query: 2   KKYLNQNVYDALIERLHPLFNDFPIVYISPSGGKDSGLLLNILLDFRDKYYPDRIG--- 58
            K Y + +V+ A + RL  +F +F  V ++PSGGKDS + L + LD        RE+G
Sbjct: 4   KHYQDADVHAATLSRLRLVFRNFERVCVAPSGGKDSSVTLQLALDVA------RELGRSP 57

Query: 59  --VFHQDFEAQYSLITKYVQETFTSLEGRKKVSLYWVCLPMATRTALSSYEMFWYPWDDK 116
              V   D E QY  T  +V E      GR  V  +WVCLP+  R A S  E +W  W+
Sbjct: 58  VDVLFIDLESQYQATIDHVSEML----GRDVRPWWVCLPLNLRNASSLEEPYWCCWEPG 113

Query: 117 TEDIWVRFMPSQDYVINLENNSITTYRYKMNQEDLAKQPGRWYKQIHGNQKTVCILGNRA 176
```

```
              E  WVRP+P Q  VI+ +       YRY+M  E+     F  W  +     + T  ++G R+
Sbjct: 114 AEADWVRPLPKQRGVIS-DPAFPPFYRYRMEFEEFVAGFNAWLAR---EEPTAFLVGIRS 169

Query: 177 SESLHRYSGFINKKYGYQKEC-----------WITKQFKDVWTAS--PLYDWSVEDIWH 222
            ESL+RY     K+     K+C            W  +   +   S  P+YDW  ED+W
Sbjct: 170 DESLNRYLAV--KKRSRAKQCAWTPPGGSAPLAWSARDRANPQAVSFFPIYDWRFEDLWR 227

Query: 223 AYYKFSYSYNELYDLPYKAGLKPSQMRVASPFQDYAVDSLNLYRIIDQETWVKLLGRVQG 282
                Y+YN LYD  Y+AG+  SQMR+  P+ D     L+L+  I+  TW K++ RV G
Sbjct: 228 CVADHGYAYNRLYDQMYRAGVPPSQMRICQPVGDDQRKGLDLFHRIEPRTWFKVRRVAG 287

Query: 283 VNFSNIYGRTKAMGYK-SLALPKGH-SWKSYTQFLLSTLPVRLRNNYVRKFNKSIDFWHK 340
             N+   Y R + +GY+  + LP     +W+ Y+QFLL ++P  LR  Y R+  + I +W +
Sbjct: 288 ANYGARYCRQRFLGYRGKGLGLPPSFGTWREYSQFLLRSMPPPLRGIYQRRIERFILWWKQ 347

Query: 341 TGGGLAEFTINELIEKGYRIARNGISNYTSFKHSRVIVLDQ-IPDDTDDIVTTKDIPSWK 399
                LA                                I+ D  IP      +  PSW+
Sbjct: 348 HDYPLA-------------------------------IWPDAGIP----ALENRRQPSWR 373

Query: 400 RMCFCILKNDHICRTMGFGLTR 421
            R+    +LK D + R++ FG ++
Sbjct: 374 RIALSLLKQD-MARSLSFGFSQ 394
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2421

A DNA sequence (GASx189) was identified in *S.pyogenes* <SEQ ID 7341> which encodes the amino acid sequence <SEQ ID 7342>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4121(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC73702 GB:AE000165 orf, hypothetical protein [Escherichia
          coli]
 Identities = 79/162 (48%), Positives = 110/162 (67%), Gaps = 1/162 (0%)

Query: 7   PVYEIKSIPIEKISPNDYNPNSVADPEMKLLYDSIKSDGYTMPIVCYYDKEEDRYSIVDG 66
           PV  +  +  ++ PNDYNPN+VADPE KLL  SI+ DG+T PIV  +  +++   IVDG
Sbjct: 46  PVDCVLWVKNSQLMPNDYNPNNVADPEKKLLQKSIRIDGFTQPIVVYNT-DKNANKIVDG 104

Query: 67  FHRYRIMLDYSDIYERESGRLPVSVIDKSLDYRMASTIRHNRARGSHDVDLMSQIVKDLH 126
           FHR+ I     S +  R  G LPV+ ++ + + R+R+TIRHNRARG H +  MS+IV++L
Sbjct: 105 FHRHEIGKGSSSLKLRLKGYLFVTCLEGTRNQRIAATIRHNRARGRHQITAMSEIVRELS 164

Query: 127 EGGRSDNWIAKHLGMDKDEILRLKQITGLASLFKDHEFNQSW 168
           + G  DN I K LGMD DE+LRLKQI GL  LF D +++++W
Sbjct: 165 QLGWDDNKIGKELGMDSDEVLRLKQINGLQELFADRQYSRAW 206
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2422

A repeated DNA sequence (GASx192R) was identified in *S.pyogenes* <SEQ ID 7343> which encodes the amino acid sequence <SEQ ID 7344>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4301(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA63509 GB:X92946 transposase [Lactococcus lactis]
 Identities = 23/36 (63%), Positives = 28/36 (76%)

Query: 1   MQDKLVTRAFNQAYNREKPKEGVIVHTDQGSQYTGA 36
           MQDKLV + F QA  +E P+ G+IVHTDQGSQYT +
Sbjct: 134 MQDKLVRDCFLQACGKEHPQPGLIVHTDQGSQYTSS 169
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2423

A DNA sequence (GASx194R) was identified in *S.pyogenes* <SEQ ID 7345> which encodes the amino acid sequence <SEQ ID 7346>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA63508 GB:X92946 hypothetical protein [Lactococcus lactis]
 Identities = 64/96 (66%), Positives = 78/96 (80%)

Query: 1  MPRKTPDKAPKLSAVKLILEEEQSVKMVSSTLEIHPNSLYQWICEYEKYGESAPGHGSA 60
          M R+ PDK PK SAVKLILEE  SVK VS  LE+H NSLY+W+QE E+YGESAPG+G+A
Sbjct: 1  MARRKPDKQPKNSAVKLILEEGYSVKEVSQELEVHANSLYRWVQEVEEYGESAPGNGTA 60

Query: 61 LRHAQFETKKLEKEHKLLQEELALLKKFQVFLKPNR 96
          L +AQ + K LEKE++ LQEEL LLKKF+VFLK ++
Sbjct: 61 LANAQHKIKILEKENRYLQEELELLKKFKVFLKRSK 96
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2424

A DNA sequence (GASx195R) was identified in *S.pyogenes* <SEQ ID 7347> which encodes the amino acid sequence <SEQ ID 7348>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-11.30   Transmembrane  179 - 195 ( 173 - 201)
   INTEGRAL    Likelihood = -8.86   Transmembrane  229 - 245 ( 224 - 254)
   INTEGRAL    Likelihood = -8.39   Transmembrane  289 - 305 ( 280 - 307)
   INTEGRAL    Likelihood = -8.23   Transmembrane  417 - 433 ( 410 - 435)
   INTEGRAL    Likelihood = -5.89   Transmembrane  324 - 340 ( 323 - 349)
   INTEGRAL    Likelihood = -4.73   Transmembrane  260 - 276 ( 256 - 278)
   INTEGRAL    Likelihood = -4.51   Transmembrane   96 - 112 (  91 - 113)
   INTEGRAL    Likelihood = -4.25   Transmembrane   24 -  40 (  20 -  43)
   INTEGRAL    Likelihood = -2.44   Transmembrane  344 - 360 ( 342 - 360)

----- Final Results -----
               bacterial membrane --- Certainty=0.5522(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB75191 GB:AL139075 putative integral membrane protein
          [Campylobacter jejuni]
 Identities = 177/430 (41%), Positives = 274/430 (63%), Gaps = 8/430 (1%)

Query: 5   IIISAIALAIGIGYRTKINIGLLAIAPSYLIATTLMGLSPKELLHFWPTSLFFTIFSVSL 64
           +IIS+I +AI +GY T+ N+G+ A+ F+Y+I     M L+PK+++ FWP S+FF IF+VSL
Sbjct: 6   LIISSIIVAIILGYITRHNVGIPAMIFAYIIGAFFMDLAPKKIIAFWPISIFFVIFAVSL 65

Query: 65  FYNVATTNGTLDVLAQHILYRTRTHPNALYMILYLIATLLSALGAGFFTTMAVCCPLAIT 124
           FYN AT NGTL+ LA H++YR   HP  L  ++++++ +++ALGAGF+T  +A     PL
Sbjct: 66  FYNFATVNGTLSKLAGHLMYRPAMIPYLLPFVIFVVSAIIAALGAGFYTVLAFMAPLTFL 125

Query: 125 LCQKADKHPLIGAQAVNWGASGGANLITSGSGIVFQGLFKQMGWE-EQAPSLGNHIFIVS 183
           LC K     + GA A+N+GA GGAN ITS SGI+F+GL +  G E  +AP+  + IF  +
Sbjct: 126 LCDKIGLSKIAGAMAINYGALGGANFITSQSGIIFRGLMENSGIRANEAPANSSIIFAFT 185

Query: 184 IIYPLIVLLLLSCYIRYSKGRTNSSLT-IDQPPVLSKVQRQTTLLMISSMVLVWLFPLLL 242
           II P++VL     +  ++  + N  ++ I +P      Q+ T +LM    +V+V +FP+L
Sbjct: 186 IILPIVVL----SFFVFNAFKNNIKISVISKPDPFDYKQKTTLIIMFMMIVVVLIPFVLN 241

Query: 243 LIFFNIAWIATYRQTFDIGFVSILMVCLALRLKLGKQEAILAKVIWAIIIMLCGMSLLMS 302
           +IFF+  I+ + +  DI  ++++ V +AL LKL  ++ ++A +FW  +IM+CG+ +L+S
Sbjct: 242 IIFFINETISYFNKKIDIAMIAMIFVAIALFLKLADEKQVVALIPWGTLIMICGVGMLIS 301

Query: 303 LAVKSGLVTLIGHLITTTIPHFWLPLFFCVIAGVMSLFSSTLSVVAPTLFPIIATISAQS 362
           +AV++G + L  L+ .  I  ++PL  C IA  MSLFSSTL VV P LFPI+ +I+A S
Sbjct: 302 IAVEAGAIKLFSDLVENEINVIFIPLIMCAIAAPMSLFSSTLGVVIPALFPIVPSIAASS 361

Query: 363 PHIDIRLLTTATIIGALSTNISPFSSAGSLIQLSLPHIFKRRSLAPKKQILLGVPISLSLA 422
             +  LL +  ++GA ++ ISPFSS GSLI  S P  +  L FK  ++  VPI     A
Sbjct: 362 GLSEALLFSCIVVGAQASAISPFSSGGSLILGSCPDKYKEKL-FKDLLIIKAVPIGFIAA 419

Query: 423 LLITWILMLL 432
           +L   I+  +
Sbjct: 420 ILATIIMSFI 429
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2425

A DNA sequence (GASx196) was identified in *S.pyogenes* <SEQ ID 7349> which encodes the amino acid sequence <SEQ ID 7350>. Analysis of this protein sequence reveals the following:

```
Possible site: 57

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0563(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC45128 GB:U65510 nicotinate-nucleotide pyrophosphorylase
            [Rhodospirillum rubrum]
 Identities = 116/277 (41%), Positives = 170/277 (60%), Gaps = 4/277 (1%)

Query: 17  LTPFQIDDTLKAALREDV-HSEDYSTNAIFDHHGQAKVSLFAKEAGVLAGLTVFQRVFTL 75
           L+PF ID+ ++ AL ED+  + D ++ A      +A     A++ G+LAGL   +  F L
Sbjct: 10  LSPFAIDEAVRRALAEDLGRAGDITSTATIFAATRAHARFVARQPGILAGLGCARSAFAL 69

Query: 76  FDTEVTPQNPHQFKDGDRLTSGDLVLEIIGSVRSLLTCERVALNFLQHLSGIASMTAAIV 135
            D  VTF   P   +DG   + +G   V E+ G+ R++L  RR ALNFL HLSGIA+ T  +
Sbjct: 70  LDDTVIFTTP--LSDGRSIAAGQTVAEVAGRARTILAARSTALNFLGHLSGIATKTRRFG 127

Query: 136 EALGDDRIKVFDTRKTTPNLRLFEKYAVRVGGGYNHRFNLSDAIMLKDNHIAAVGSVQKA 195
           +A+   R ++  TRKTTP LR  EKYAVR GGG NHRF L DA+++KDNHIA  G V  A
Sbjct: 128 DAIAHTRARLTCTRKTTPGLRGLEKYAVRCGGGSNHRFGLDDAVLIKDNHIAVAGGVSAA 187

Query: 196 IAQARAYAPFVKMVEVESL-AAAEEAAAAGVDIIMLDNMSLKQIRQAITLIAGRSRIE 254
           +++ARA     +E+EV++L   AE  A  G ++++LDNM     + +A+ ++AGR   E
Sbjct: 188 LSRARAGVGHMVRIEIEVDTLEQLAEVLAVGGAEVVLLDNMDAFTLTRAVDMVAGRLVTE 247

Query: 255 CSGNIDMTTISRFRGLAIDYVSSGSLTHSAKSLDFSM 291
            SG + + TI+      +DY+S G+LTHS  +LD  +
Sbjct: 248 ASGGVSLDTIAALAESGVDYISVGALTHSVTTLDIGL 284
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2426

A DNA sequence (GASx199) was identified in *S.pyogenes* <SEQ ID 7351> which encodes the amino acid sequence <SEQ ID 7352>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1649(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2427

A DNA sequence (GASx201) was identified in *S.pyogenes* <SEQ ID 7353> which encodes the amino acid sequence <SEQ ID 7354>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2428

A DNA sequence (GASx203) was identified in *S.pyogenes* <SEQ ID 7355> which encodes the amino acid sequence <SEQ ID 7356>. Analysis of this protein sequence reveals the following:

```
Possible site: 37

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2429

A DNA sequence (GASx210) was identified in *S.pyogenes* <SEQ ID 7357> which encodes the amino acid sequence <SEQ ID 7358>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2430

A DNA sequence (GASx211) was identified in *S.pyogenes* <SEQ ID 7359> which encodes the amino acid sequence <SEQ ID 7360>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
              bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2431

A DNA sequence (GASx213) was identified in *S.pyogenes* <SEQ ID 7361> which encodes the amino acid sequence <SEQ ID 7362>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4430(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2432

A DNA sequence (GASx219) was identified in *S.pyogenes* <SEQ ID 7363> which encodes the amino acid sequence <SEQ ID 7364>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2433

A DNA sequence (GASx220) was identified in *S.pyogenes* <SEQ ID 7365> which encodes the amino acid sequence <SEQ ID 7366>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0530(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2434

A DNA sequence (GASx231R) was identified in *S.pyogenes* <SEQ ID 7367> which encodes the amino acid sequence <SEQ ID 7368>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2435

A DNA sequence (GASx237) was identified in *S.pyogenes* <SEQ ID 7369> which encodes the amino acid sequence <SEQ ID 7370>. Analysis of this protein sequence reveals the following:

```
Possible site: 52

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4961(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB49143 GB:AJ248283 hypothetical protein [Pyrococcus abyssi]
 Identities = 79/229 (34%), Positives = 131/229 (56%), Gaps = 11/229 (4%)

Query: 18  MRFTIDQNMQPFLVRIDLSHGGSVYLQQGSMVYHTRNVTLNTKINGKGSGLGKLVGAIGR 77
           M + I+    F L+E++L  G +V  + G+MVY     V++ TK  G        L+GA+ R
Sbjct: 1   MEYRIEHRPSPSLLEVNLRSGEAVQAEAGAMVYMDPTVSIETKARGG------LLGALKR 54

Query: 78  SMVSGESMFITQAMSNGDGKLALAPNTPGQIVALELGEKQYRLNDGAFLALDGSAQYKME 137
           S++ GES F+   +  G G++  AP  PG I++LEL    Y    GAFL          ++
Sbjct: 55  SVLSGESFFMN--VFRGPGRVGFAPGYPGDIISLELNGTLYA-QSGAFLVASEGIDIDVK 111

Query: 138 RQNIGKALPGGQGGLFVMTTRGLGTLLANSPGSIKKITLDGSIWTIDNAHVVAWSRELDY 197
               GK +PG +G +P++  +G G +  +S+G+I+KITL G ++ +D  H+VA++  +D+
Sbjct: 112 FGG-GKYIPGRHG-VFLELKGKGIVFLSSYGAIEKITLRGESVIVDTGHMVAPTEGIDF 169

Query: 198 DIHLENGFMQSIGTGEGVVMTPRGHGEIYIQSLNLEQFAGTLKRYLPTS 246
            I   G   ++ +GEG+V  F GHG++YIQ+ +L+ F   + +LP S
Sbjct: 170 RIKIGGLKATLPSGEGLVFRFSGHGDVYIQTRSLDGFLSWILPHLPKS 218
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2436

A DNA sequence (GASx240R) was identified in *S.pyogenes* <SEQ ID 7371> which encodes the amino acid sequence <SEQ ID 7372>. Analysis of this protein sequence reveals the following:

```
Possible site: 35

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2745(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2437

A DNA sequence (GASx241) was identified in *S.pyogenes* <SEQ ID 7373> which encodes the amino acid sequence <SEQ ID 7374>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-10.14   Transmembrane  196 - 212 ( 187 - 215)
   INTEGRAL    Likelihood = -8.01   Transmembrane  160 - 176 ( 156 - 179)
   INTEGRAL    Likelihood = -5.89   Transmembrane  116 - 132 ( 110 - 134)
   INTEGRAL    Likelihood = -4.57   Transmembrane   74 -  90 (  73 -  97)
   INTEGRAL    Likelihood = -2.66   Transmembrane   51 -  67 (  50 -  68)
   INTEGRAL    Likelihood = -2.60   Transmembrane    8 -  24 (   7 -  27)
   INTEGRAL    Likelihood = -1.28   Transmembrane  344 - 360 ( 344 - 360)
   INTEGRAL    Likelihood = -0.22   Transmembrane   30 -  46 (  30 -  46)
```

```
----- Final Results -----
                bacterial membrane --- Certainty=0.5055(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAC10175 GB:AJ278302 histidine kinase [Streptococcus pneumoniae]
 Identities = 136/449 (30%), Positives = 234/449 (51%), Gaps = 26/449 (5%)

Query: 8   FLLLSIIVYYMTKIYIFSFLSDITLP---VWKQLTI-LALALFFNQFFYLS-----PLLI 58
           ++LL  +V +  KI IF  +  I+L     ++K     + LA+ F    Y+          +
Sbjct: 5   WILLYTLVTHGLKIVIFFKVDGISLTFSRIFKAFLFKILLAAVVFGNLGYMVGNVYLSYFN 64

Query: 59  DPL----LFLVVLRQETKQLFSLKALFLAVAPSVLVDLLSRFMGTIVIPYLPLSSGIYLG 114
           +PL     L  ++LR+  K+L      LF  + P +LV+L  R +   V+P  FL  G
Sbjct: 65  EPLYGIGLSFLLLRELPKKLL----LFYGLFPMILVNLFYRGVSYFVLP--FLGQGQVYD 118

Query: 115 HIIFDLLAYLLDIPPSFAIINYMIGKDYKMIC-QSGYSKRSHNFYQTLLMFVLVYYVDIFV 173
            20
           Sbjct: 119 DYSFIWLC-IIIFNFPISLAFLKWLDYDPTSLRKGILDKDPQKSLTQINWIMGAYYLVIQ 177

Query: 174 ILGFTDPFLHFHHSLFVPTPYKLLFLMPILLLVYLLSYFNHSSKEYLKNELRREQQAYMT 233
                L + +    +   +   T   L+ + ++L  + ++   +   K+ L   L +EQ
Sbjct: 178 NLSYFE----YEQGIQSTTVRHLILVFYLLFFMGIIKKLDTYLKDKLHERLNQEQDLRYR 233

Query: 234 NLEIYGKHLEKLYRDVRAPQSDYLSRIERLGQAIKSESITQIQDIYAQTVHEANDYWDDK 293
             +E Y +H+E+LY++VR+F+ DY + +  L   I+ E + QI++IY   + ++++   D
Sbjct: 234 EMERYSRHIEELYKEVRSFRHDYTNLLTSLRLGIEEEDMSQIKEIYDSVLKDSSEKLQDM 293

Query: 294 HYNISKLREKINISSIKGLLGAKIISAEKSGDILNVEVDDNIKETYIPELDLLLLMSIFCD 353
             Y++ +L  +   ++KSLL+ K I A    I  NVEVP+ I+   + LD L ++SI CD
Sbjct: 294 KYDLGRLNVKDRAALKSLLAGKFIKARDKNIVFNVEVPEEIQVEGVSLLDFLTVVSILCD 353

Query: 354 NAIEAALEAQQPHMSIAYFLLGDYQMFVVTNTTKKK-VDINKIPEEGYSSKGSERGIGLS 412
           NAIEA++EA QPH+SIA+F  G  + F++ N+ K++ +DI++IF  G SSKG ERG+GL
Sbjct: 354 NAIEASVEACQPHVSIAFFKNGAQETFIIENSIKEEGIDISEIPSFGASSKGEERGVGLY 413

Query: 413 NAQRILKKYPYLSLRTKSFDKEPSQTLTM 441
            +I++ +P  SL T   D  F Q LT+
Sbjct: 414 TVMKIVESHPNTSLNTTCQDHVFRQVLTV 442
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2438

A DNA sequence (GASx242R) was identified in *S.pyogenes* <SEQ ID 7375> which encodes the amino acid sequence <SEQ ID 7376>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.4165(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2439

A DNA sequence (GASx243) was identified in *S.pyogenes* <SEQ ID 7377> which encodes the amino acid sequence <SEQ ID 7378>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-11.09   Transmembrane  188 - 204 ( 182 - 208)
   INTEGRAL    Likelihood = -7.17   Transmembrane   52 -  68 (  47 -  69)
   INTEGRAL    Likelihood = -4.73   Transmembrane  119 - 135 ( 114 - 142)
   INTEGRAL    Likelihood = -4.62   Transmembrane   83 -  99 (  77 - 107)
   INTEGRAL    Likelihood = -1.86   Transmembrane  328 - 344 ( 328 - 345)
   INTEGRAL    Likelihood = -1.65   Transmembrane    7 -  23 (   6 -  23)
   INTEGRAL    Likelihood = -0.22   Transmembrane   35 -  51 (  35 -  51)

----- Final Results -----
              bacterial membrane --- Certainty=0.5437(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAC10175 GB:AJ278302 histidine kinase [Streptococcus pneumoniae]
 Identities = 123/438 (28%), Positives = 229/438 (52%), Gaps = 49/438 (11%)

Query: 20  VIFAKVSAIKLSWKRVS-------IIGISFVIANMIFDKVIL---IDGLFFIVSLL--- 66
           VIF KV  I L+++R+         ++ + F +   +  V L   ++ L+ I +S L
Sbjct: 19  VIFFKVDGISLTPERIFKAFLFKILLSVVFGMLGYMVGNVYLSYFMEPLYGIGLSFLLLR 78

Query: 67  SAPKKKLFEHMFNGFFTILIVELLFRVIGSFFLPAVLGFSIGQINNNLKLLELCYLFVLP 126
             PKK L   +F G F +++V L +R +  F LP +     GQ+ ++   + LC + +
Sbjct: 79  ELPKKLL---LFYGLFPMILVNLFYRGVSYFVLPFL---GGGQVYDDYSFIWLC-IIIFN 131

Query: 127 IFYLFSYIFSDIL---SLIRFISEDKMKMVFWWMTAMPSYYFFAHFLVTVQSGFPLALYF 183
            F   +++  +D    SL + I +   +K +  +N  M +YY     L        YF
Sbjct: 132 FFISLAFLKWLDYDPTSLRKGILDKDPQKSLTQINWIMGAYYLVIQNLS---------YF 182

Query: 184 QY---------RSILVFIYLAIFIWVIVKLDRFAKDQLSQKLTQAQNERIAYLENYNQSI 234
           +Y         R +++   YL  F+ +I KLD + KD+L ++L Q Q+ R   +E Y++ I
Sbjct: 183 EYEQGIQSTTVRHLILVFYLLFPMGIIKKLDTYLKDKLHERLNQEQDLRYREMERYSRHI 242

Query: 235 EQLYREIRTVKHDSENILISLKDSIDSGDIDLITRVYDTVIQQSATSMMRTNYEISSLDN 294
           E+LY+E+R+ +HD  N+L SL+  I+  D++ I  +YD+V++ S+  +    Y++  L N
Sbjct: 243 EELYKEVRSFRHDYTNLLTSLRLGIEKEDMKQIKSTYDSVIKDSSSKLQDNKYDLGRLVN 302

Query: 295 IKEAVIRSIMNSKLLHAQVLGIELVYIKIPDVIDHLPIKLIDLIVLFTGLVDNAIETAKGS 354
           +++  ++S++  K ++A+   I   +E+P+ I  + L+D + + + L DNAIE +  +
Sbjct: 303 VRDRALKSLLAGKFIKARDKNIVFNVEVPEEIQVEGVSLLDFLTVVSILCTNAIEASVEA 362

Query: 355 RRPFLSIAYFKQDNKQLPIIENSTKTNRVDIAKRFDAQQQNSAH--------FLTVLDSY 406
             +P +SIA+FK    + FIIKNS K   +DI++ F     +             + +++S+
Sbjct: 363 CQPHVSIAFFKNGAQETPIIENSIKERGIDISEIFSPGASSKGEERGVGLYTVMKIVESH 422

Query: 407 PQITLSTKSDHYRLRQLL 424
           P  +L+T   +  RQ+L
Sbjct: 423 PNTSLNTTCQDHVFRQVL 440
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2440

A DNA sequence (GASx248) was identified in *S.pyogenes* <SEQ ID 7379> which encodes the amino acid sequence <SEQ ID 7380>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.5665(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2441

A DNA sequence (GASx255) was identified in *S.pyogenes* <SEQ ID 7381> which encodes the amino acid sequence <SEQ ID 7382>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1437(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2442

A DNA sequence (GASx270R) was identified in *S.pyogenes* <SEQ ID 7383> which encodes the amino acid sequence <SEQ ID 7384>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -5.89   Transmembrane   20 -  36 (  17 -  36)

----- Final Results -----
             bacterial membrane --- Certainty=0.3357(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2443

A DNA sequence (GASx272) was identified in *S.pyogenes* <SEQ ID 7385> which encodes the amino acid sequence <SEQ ID 7386>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2488(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB11887 GB:Z99104 ribosomal protein S7 (BS7) [Bacillus subtilis]
 Identities = 117/156 (75%), Positives = 139/156 (89%)

Query: 1   MSRKNQAPKREVLPDPLYNSKLVTRLINRVMLDGKRGTAATIVYDAFNAIKEATGNDALE 60
           M RK    KR+VLPDP+YNSK+V+RLIN++M+DGK+G   TI+Y +F+ IKE TGNDA+E
Sbjct: 1   MPRKGPVAKRDVLPDPIYNSKLVSRLINKMMIDGKKGKPQTILYKSFDIIKERTGNDAME 60

Query: 61  VFETAMDNIMPVLEVRARRVGGSNYQVPVEVRPERRTTLGLRWLVNASRARGEHTMKDRL 120
           VFE A+ NIMPVLEV+ARRVGG+NYQVEVEVRPERRTTLGLRWLVN +R RGE TM++RL
Sbjct: 61  VFEQALKNIMPVLEVKARRVGGANYQVEVEVRPERRTTLGLRWLVNYARLRGEKTMEERL 120

Query: 121 AKEIMDAANNTGASVKKREDTHKMAEANRAFAHFRW 156
           A EI+DAANNTGA+VKKREDTHKMAEAN+AFAH+RW
Sbjct: 121 ANEILDAANNTGAAVKKREDTHKMAEANKAFAHYRW 156
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2444

A DNA sequence (GASx274) was identified in *S.pyogenes* <SEQ ID 7387> which encodes the amino acid sequence <SEQ ID 7388>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

A related sequence was also identified in GAS <SEQ ID 9095> which encodes the amino acid sequence <SEQ ID 9096>. Analysis of this protein sequence reveals the following:

```
     Possible cleavage site: 52
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty= 0.291(Affirmative) < succ>
             bacterial membrane --- Certainty= 0.000(Not Clear) < succ>
              bacterial outside --- Certainty= 0.000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2445

A DNA sequence (GASx275R) was identified in *S.pyogenes* <SEQ ID 7389> which encodes the amino acid sequence <SEQ ID 7390>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.5664(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2446

A DNA sequence (GASx283) was identified in *S.pyogenes* <SEQ ID 7391> which encodes the amino acid sequence <SEQ ID 7392>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0724(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2447

A DNA sequence (GASx298) was identified in *S.pyogenes* <SEQ ID 7393> which encodes the amino acid sequence <SEQ ID 7394>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2840(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2448

A DNA sequence (GASx300) was identified in *S.pyogenes* <SEQ ID 7395> which encodes the amino acid sequence <SEQ ID 7396>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -1.91   Transmembrane    4 -  20 (   4 -  20)

----- Final Results -----
               bacterial membrane --- Certainty=0.1765(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2449

A DNA sequence (GASx301) was identified in *S.pyogenes* <SEQ ID 7397> which encodes the amino acid sequence <SEQ ID 7398>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4884(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2450

A repeated DNA sequence (GASx302) was identified in *S.pyogenes* <SEQ ID 7399> which encodes the amino acid sequence <SEQ ID 7400>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
```

```
bacterial cytoplasm --- Certainty=0.2581(Affirmative) < succ>
 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2451

A DNA sequence (GASx316) was identified in *S.pyogenes* <SEQ ID 7401> which encodes the amino acid sequence <SEQ ID 7402>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.80   Transmembrane   23 -  39 (  22 -  39)

----- Final Results -----
               bacterial membrane --- Certainty=0.1319(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2452

A DNA sequence (GASx323R) was identified in *S.pyogenes* <SEQ ID 7403> which encodes the amino acid sequence <SEQ ID 7404>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.0005(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2453

A DNA sequence (GASx334) was identified in *S.pyogenes* <SEQ ID 7405> which encodes the amino acid sequence <SEQ ID 7406>. Analysis of this protein sequence reveals the following:

```
Possible site: 17
```

```
>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2454

A DNA sequence (GASx336) was identified in *S.pyogenes* <SEQ ID 7407> which encodes the amino acid sequence <SEQ ID 7408>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.3379(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2455

A DNA sequence (GASx361R) was identified in *S.pyogenes* <SEQ ID 7409> which encodes the amino acid sequence <SEQ ID 7410>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.2807(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2456

A DNA sequence (GASx387) was identified in *S.pyogenes* <SEQ ID 7411> which encodes the amino acid sequence <SEQ ID 7412>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2740(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2457

A DNA sequence (GASx389) was identified in *S.pyogenes* <SEQ ID 7413> which encodes the amino acid sequence <SEQ ID 7414>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0744(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2458

A DNA sequence (GASx392) was identified in *S.pyogenes* <SEQ ID 7415> which encodes the amino acid sequence <SEQ ID 7416>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2162(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2459

A DNA sequence (GASx393R) was identified in *S.pyogenes* <SEQ ID 7417> which encodes the amino acid sequence <SEQ ID 7418>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2520(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2460

A DNA sequence (GASx395) was identified in *S.pyogenes* <SEQ ID 7419> which encodes the amino acid sequence <SEQ ID 7420>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2590(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2461

A DNA sequence (GASx396) was identified in *S.pyogenes* <SEQ ID 7421> which encodes the amino acid sequence <SEQ ID 7422>. Analysis of this protein sequence reveals the following:

```
Possible site: 41

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB13373 GB:Z99111 similar to hypothetical proteins [Bacillus subtilis]
 Identities = 23/88 (26%), Positives = 52/88 (58%)
```

```
Query: 4   KQERIGLVVYLYYNKDARKLSKFGDLYYHSKRSRYLIIYINKNDLDTKLEEMRRLKCVKD 63
           +  R G+VVYL+  + ++ L KFG+++Y SKR +Y+++Y + + ++  ++++      VK
Sbjct: 2   ENRRQGMVVYLHSLKQGKMLRKFGNVHYVSKRLKYVVLYCDMDQIEKTMDKIASYSFVKK 61

Query: 64  IRPSAFDDIDRQFVGNLHRDETNNHQKG 91
           + PS    +  +F   L + +  +++ G
Sbjct: 62  VEPSYKPFLKLEFESKLDKAKEYDYKIG 89
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2462

A DNA sequence (GASx400) was identified in *S.pyogenes* <SEQ ID 7423> which encodes the amino acid sequence <SEQ ID 7424>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2010(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2463

A DNA sequence (GASx401) was identified in *S.pyogenes* <SEQ ID 7425> which encodes the amino acid sequence <SEQ ID 7426>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1176(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2464

A DNA sequence (GASx402) was identified in *S.pyogenes* <SEQ ID 7427> which encodes the amino acid sequence <SEQ ID 7428>. Analysis of this protein sequence reveals the following:

```
Possible site: 16
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.2938(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2465

A DNA sequence (GASx403R) was identified in *S.pyogenes* <SEQ ID 7429> which encodes the amino acid sequence <SEQ ID 7430>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                 bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2466

A DNA sequence (GASx406) was identified in *S.pyogenes* <SEQ ID 7431> which encodes the amino acid sequence <SEQ ID 7432>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-12.26   Transmembrane   15 -  31 (   4 -  36)
   INTEGRAL    Likelihood = -6.64   Transmembrane   96 - 112 (  94 - 115)

----- Final Results -----
                bacterial membrane --- Certainty=0.5904(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2467

A DNA sequence (GASx408R) was identified in *S.pyogenes* <SEQ ID 7433> which encodes the amino acid sequence <SEQ ID 7434>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -2.23   Transmembrane   17 -  33 ( 15 -  34)
   INTEGRAL    Likelihood = -0.85   Transmembrane   38 -  54 ( 38 -  54)

----- Final Results -----
               bacterial membrane --- Certainty=0.1893(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2468

A DNA sequence (GASx412) was identified in *S.pyogenes* <SEQ ID 7435> which encodes the amino acid sequence <SEQ ID 7436>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -6.53   Transmembrane    5 -  21 (  4 -  23)

----- Final Results -----
               bacterial membrane --- Certainty=0.3612(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2469

A DNA sequence (GASx413) was identified in *S.pyogenes* <SEQ ID 7437> which encodes the amino acid sequence <SEQ ID 7438>. Analysis of this protein sequence reveals the following:

```
Possible site: 56

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3422(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA68903 GB:Y07622 lactate oxidase [Streptococcus iniae]
 Identities = 328/392 (83%), Positives = 359/392 (90%), Gaps = 4/392 (1%)

Query: 3   MAQKTVITEETTDFVMDFKTSSABGNVDFINVFDLEKMAQQVIPKGAFGYIASGAGDTFT 62
            M  K+ +   TT   ++PKTSSABG+VDF+NVFDLEKMAQ+VIPKGAFGYIASGAGDTFT
Sbjct: 1   MSNKSEMINATT---IEPKTSSABGSVDFVNVFDLEKMAQKVIPKGAFGYIASGAGDTFT 57

Query: 63  LHENIRSFMHKLIVPHSLKGVENPSTEITFDGDYLTSPLILAPVAAHKLANEQGEVASAK 122
            L+ENIRSFMHKLI  PH LKGVENPSTEITF GD L SP+HLAPVAAHKLANEQGE+ASAK
Sbjct: 58  LNENIRSFMHKLI-PHGLKGVENPSTEITFIGDKLASPIILAPVAAHKLANEQGEIASAK 116

Query: 123 GLKEFGSIYTTSSYSTTDLPEISAALGGTPHWFQFYYSKDDGINRNIMDRVKAQGCKAIV 182
            G+KEFG+IYTTSSYSTTDLPEIS  LG +PHWFQFYYSKDDGINR+IMDR+KA+G K+IV
Sbjct: 117 GVKEFGTIYTTSSYSTTDLPEISQTLGDSPHWFQFYYSKDDGINRHIMDRLKASGVKSIV 176

Query: 183 LTALATVGGNREVDRRNGFVFFVGMPIVQEYLPDGAGKTMDYVYKSAKQALTSKDIEYIA 242
            LT LATVGGNREVD+RNGFVFFVGMPIVQEYLP+GAGKTMDYVYK+ KQAL+ KD+EYIA
Sbjct: 177 LTVLATVGGNREVDKRNGFVFFVGMPIVQEYLPNGAGKTMDYVYKTKQALSPKDVEYIA 236

Query: 243 TYSGLPVYVKGPQCAEDTLRALDAGASGIWVTNHGGRQLDGGPAAFDSLQEVAEAVDQKV 302
             YSGLPVYVKGPQCAED  RAL+AGASGIWVTNHGGRQLDGGPAAFDSLQEVAE+VD++V
Sbjct: 237 QYSGLPVYVKGPQCAEDAFRALEAGASGIWVTNHGGRQLDGGPAAFDSLQEVAESVDRRV 296

Query: 303 PIVFDSGIRRGQHIFKALASGADLVALGRPAIYGLAMGGSIGTRQVFEKLNDELKMVMQL 362
            PIVFDSG+RRGQH+FKALASGADLVALGRP IYGLAMGGSVGTRQVFEKINDELKMVMQL
Sbjct: 297 PIVFDSGVRRGQHVFKALASGADLVALGRFVIYGLAMGGSVGTRQVFEKINDELKMVMQL 356

Query: 363 AGTQTIQDVKAFNLRHNEYDSSIPFDQNALRL 394
            AGTQTI  DVK F LRHNEYDSSIPF     ++
Sbjct: 357 AGTQTIDDVKHFKLRHNEYDSSIPFSPKCFKI 388
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2470

A DNA sequence (GASx414) was identified in *S.pyogenes* <SEQ ID 7439> which encodes the amino acid sequence <SEQ ID 7440>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0682(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2471

A DNA sequence (GASx417R) was identified in *S.pyogenes* <SEQ ID 7441> which encodes the amino acid sequence <SEQ ID 7442>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence
```

```
----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1765(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2472

A DNA sequence (GASx418) was identified in *S.pyogenes* <SEQ ID 7443> which encodes the amino acid sequence <SEQ ID 7444>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2532(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2473

A DNA sequence (GASx419) was identified in *S.pyogenes* <SEQ ID 7445> which encodes the amino acid sequence <SEQ ID 7446>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3082(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2474

A DNA sequence (GASx423) was identified in *S.pyogenes* <SEQ ID 7447> which encodes the amino acid sequence <SEQ ID 7448>. Analysis of this protein sequence reveals the following:

```
Possible site: 52
```

```
>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -2.18   Transmembrane   14 -  30 (  13 -  31)

----- Final Results -----
                bacterial membrane --- Certainty=0.1871(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2475

A DNA sequence (GASx427R) was identified in *S.pyogenes* <SEQ ID 7449> which encodes the amino acid sequence <SEQ ID 7450>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -1.17   Transmembrane   13 -  29 (  10 -  29)

----- Final Results -----
                bacterial membrane --- Certainty=0.1468(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

A related sequence was also identified in GAS <SEQ ID 9105> which encodes the amino acid sequence <SEQ ID 9106>. Analysis of this protein sequence reveals the following:

```
     Possible site: 20
>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -1.17   Transmembrane   8 -  24

----- Final Results -----
                bacterial membrane --- Certainty=0.1470(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA26616 GB:M63917 epidermal cell differentiation inhibitor
          [Staphylococcus aureus]
 Identities = 58/195 (29%), Positives = 106/195 (53%), Gaps = 13/195 (6%)

Query: 67  RWGKGLI----YPRAEQEAMAAYTCQQAGPINTSLDKAKGELSQLTPELRDQVAQLDAAT 122
           +WG LI    Y   ++ A+  YT + +  IN  L  A G++++L     +D+V +LD++
Sbjct: 49  KWGNKLIKQAKYSSDDKIALYEYT-KDSSKINGFLKLAGGDINKLDSTTQDKVRRLDSSI 107

Query: 123 HRLVIPWNIVVYRYVYETFLRDI-GVSHADLTSYYR--NHQPDPHILCKIK--LGTR-YT 176
            +   P ++ VYR +   +L  I G ++ DL     +  N Q+D +++ K+    + +R Y
Sbjct: 108 SKSTTPESVYVYRLLNLDYITSIVGFTNEDLYKLQQTNNGQYDENLVRKLNNVMNSRIYR 167

Query: 177 KHSFMSTTALKNGAMTHRPVEVRICVKTGANAAFV--EPYSAVPSEVELLFPRGCQLEVV 234
           + + ST  +   A+  RP+E+R+ + KG KAA++  +  +A   + E+L PRG +  V
Sbjct: 168 EDGYSSTQLVSGAAVGGRPIELRLELPKGTKAAYLNSKDLTAYYGQQEVLLPRGTEYAVG 227

Query: 235 GAYVSQDQKKLHIEA 249
               +S D+KK+ I A
```

```
Sbjct: 228 SVELSNDKKKIIITA 242
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2476

A DNA sequence (GASx428) was identified in *S.pyogenes* <SEQ ID 7451> which encodes the amino acid sequence <SEQ ID 7452>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3817(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2477

A DNA sequence (GASx429) was identified in *S.pyogenes* <SEQ ID 7453> which encodes the amino acid sequence <SEQ ID 7454>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
              bacterial outside --- Certainty=0.3000(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2478

A DNA sequence (GASx431) was identified in *S.pyogenes* <SEQ ID 7455> which encodes the amino acid sequence <SEQ ID 7456>. Analysis of this protein sequence reveals the following:

```
Possible site: 43

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL     Likelihood = -8.60   Transmembrane   68 -  84 (  66 -  90)
   INTEGRAL     Likelihood = -6.85   Transmembrane   22 -  38 (  16 -  42)
   INTEGRAL     Likelihood = -3.29   Transmembrane   44 -  60 (  43 -  61)

----- Final Results -----
            bacterial membrane --- Certainty=0.4439(Affirmative) < succ>
```

```
          bacterial outside --- Certainty=0.0000(Not Clear) < succ>
        bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2479

A DNA sequence (GASx432R) was identified in *S.pyogenes* <SEQ ID 7457> which encodes the amino acid sequence <SEQ ID 7458>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
          bacterial outside --- Certainty=0.3000(Affirmative) < succ>
        bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
      bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2480

A DNA sequence (GASx434) was identified in *S.pyogenes* <SEQ ID 7459> which encodes the amino acid sequence <SEQ ID 7460>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
          bacterial outside --- Certainty=0.3000(Affirmative) < succ>
        bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
      bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2481

A DNA sequence (GASx435R) was identified in *S.pyogenes* <SEQ ID 7461> which encodes the amino acid sequence <SEQ ID 7462>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -2.50   Transmembrane    4 -  20 (   3 -  21)
```

```
----- Final Results -----
                bacterial membrane --- Certainty=0.1999(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB59092 GB:M97157 pyrogenic exotoxin C [Streptococcus pyogenes]
 Identities = 110/229 (48%), Positives = 150/229 (65%), Gaps = 4/229 (1%)

Query: 4   IIKTIILVIIIFHGYGS--VKSDSE-NIKDVKLQLNYAYEIIPVDYTNCNIDYLTTHDPY 60
           IIK + ++ +I     S  +KSDS+ +I +VK  L YAY I P DY +C +++ TTH
Sbjct: 6   IIKIVFIITVILISTISPIIKSDSKKDISNVKSDLLYAYTITPYDYKDCRVNFSTTHTLN 65

Query: 61  IDISSYKKKNPSVDSEVESYITKFTKMQKVNIPGLPYIPTRYDVYYIYGGVTPSVNSNS 120
           ID   Y+ K++ + SE+     + KF ++  V++PGL YI    YIYGG+TP+ N N
Sbjct: 66  IDTQKYRGKDYYISSEMSYEASQKFKKDDHVDVPGLPYILNSHTGEYIYGGITPAQN-NK 124

Query: 121 ENSKIVGNLLIDGVQQKTLINPIKIDKPIFTIQEFDFKIRQYLMQTYKIYDPNSPYIKGQ 180
            N K++GNL I G  Q+ L N I ++K I T QE DFKIR+YLM  YKIYD  SPY+ G+
Sbjct: 125 VNHKLLGNLPISGESQQNLNNKIILEKDIVTFQEIDFKIRKYLMDNYKIYDATSPYVSGR 184

Query: 181 LEIAINGNKHESPNLYDATSSSTRSDIFKKYKDNKTINMKDFSHFDIYL 229
           +EI     KHE  +L+D+ +  TRSDIF KYKDN+ IINMKNFSHFDIYL
Sbjct: 185 IEIGTKDGKHEQIDLFDSPNEGTRSDIFAKYKDNRIINMKNFSHFDIYL 233
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2482

A DNA sequence (GASx436R) was identified in *S.pyogenes* <SEQ ID 7463> which encodes the amino acid sequence <SEQ ID 7464>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2483

A DNA sequence (GASx446) was identified in *S.pyogenes* <SEQ ID 7465> which encodes the amino acid sequence <SEQ ID 7466>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
```

```
bacterial outside --- Certainty=0.3000(Affirmative) < succ>
bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2484

A DNA sequence (GASx449) was identified in *S.pyogenes* <SEQ ID 7467> which encodes the amino acid sequence <SEQ ID 7468>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -3.82   Transmembrane    3 -  19 (   1 -  20)

----- Final Results -----
               bacterial membrane --- Certainty=0.2529(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2485

A DNA sequence (GASx450R) was identified in *S.pyogenes* <SEQ ID 7469> which encodes the amino acid sequence <SEQ ID 7470>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -1.44   Transmembrane   21 -  37 (  19 -  37)

----- Final Results -----
               bacterial membrane --- Certainty=0.1574(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2486

A DNA sequence (GASx457R) was identified in *S.pyogenes* <SEQ ID 7471> which encodes the amino acid sequence <SEQ ID 7472>. Analysis of this protein sequence reveals the following:

```
Possible site: 19
```

```
>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-15.34   Transmembrane   64 -  80 (  57 -  86)
   INTEGRAL    Likelihood =-13.43   Transmembrane   97 - 113 (  91 - 116)
   INTEGRAL    Likelihood = -5.57   Transmembrane   38 -  54 (  32 -  56)

----- Final Results -----
               bacterial membrane --- Certainty=0.7135(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2487

A DNA sequence (GASx476R) was identified in *S.pyogenes* <SEQ ID 7473> which encodes the amino acid sequence <SEQ ID 7474>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3013(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2488

A DNA sequence (GASx477) was identified in *S.pyogenes* <SEQ ID 7475> which encodes the amino acid sequence <SEQ ID 7476>. Analysis of this protein sequence reveals the following:

```
Possible site: 57

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1022(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAC03521 GB:AJ276410 BlpJ protein [Streptococcus pneumoniae]
 Identities = 47/77 (61%), Positives = 59/77 (76%)

Query: 1  MIKFAEEIQKELFHIIGGVSATDCKNHLIGGITSGALAGGVGAGMATLGVGGVAGAPAG 60
          M+   E +  E L  + GGVS+TDC+N LI G+T+G I GG GAG+ATLGV G+AGAP G
Sbjct: 5  MLSQLESVMDTEMLAKVSGGVSSTDCQNALITGVTTGIITGGTGAGLATLGVAGLAGAPVG 64
```

```
Query: 61 AHVGAIAGGLTCVGGML 77
          AH+GAI GGLTC+GGM+
Sbjct: 65 AHIGAIGGGLTCLGGMV 81
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2489

A DNA sequence (GASx478) was identified in *S.pyogenes* <SEQ ID 7477> which encodes the amino acid sequence <SEQ ID 7478>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -2.07   Transmembrane   42 -  58 (  41 -  58)
   INTEGRAL    Likelihood = -1.59   Transmembrane   22 -  38 (  22 -  38)

----- Final Results -----
               bacterial membrane --- Certainty=0.1829(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAC03520 GB:AJ276410 BlpI protein [Streptococcus pneumoniae]
 Identities = 35/56 (62%), Positives = 44/56 (78%)

Query: 1  MDNFLELQFEELVNISGGKGNIGSAIGGCLGGMLIAAAGGPITGGAAAPVCVASGI 56
          M+ F  +  EEL  +SGG+GN+GSAIGGC+G +L+AAA GPITGGAA  +CV SGI
Sbjct: 6  MEQFSVMDNEELEIVSGGRGNLGSAIGGCIGAVLLAAATGPITGGAATLICVGSGI 61
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2490

A DNA sequence (GASx482) was identified in *S.pyogenes* <SEQ ID 7479> which encodes the amino acid sequence <SEQ ID 7480>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -0.43   Transmembrane   61 -  77 (  61 -  79)

----- Final Results -----
               bacterial membrane --- Certainty=0.1171(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAC03524 GB:AJ276410 BlpM protein [Streptococcus pneumoniae]
 Identities = 22/52 (42%), Positives = 30/52 (57%)

Query: 29 MEIKKLETPHQMTIEKLAKVEGGKNNWQANVSGVIAAGSAGAAIGFPVCGVA 80
          M+ K +E FH+M I  L+ +EGGKNNWQ NV      A   G  +G  +C  +
Sbjct: 1  MDTKIMEQFHEMDITMLSSIEGGKNNWQTNVLEGGGAAFGGWGLGTAICAAS 52
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2491

A DNA sequence (GASx483) was identified in *S.pyogenes* <SEQ ID 7481> which encodes the amino acid sequence <SEQ ID 7482>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1832(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2492

A DNA sequence (GASx484) was identified in *S.pyogenes* <SEQ ID 7483> which encodes the amino acid sequence <SEQ ID 7484>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
              bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2493

A DNA sequence (GASx485) was identified in *S.pyogenes* <SEQ ID 7485> which encodes the amino acid sequence <SEQ ID 7486>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1037(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2494

A DNA sequence (GASx487) was identified in *S.pyogenes* <SEQ ID 7487> which encodes the amino acid sequence <SEQ ID 7488>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1086(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2495

A DNA sequence (GASx488) was identified in *S.pyogenes* <SEQ ID 7489> which encodes the amino acid sequence <SEQ ID 7490>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2176(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2496

A DNA sequence (GASx489R) was identified in *S.pyogenes* <SEQ ID 7491> which encodes the amino acid sequence <SEQ ID 7492>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2497

A DNA sequence (GASx490) was identified in *S.pyogenes* <SEQ ID 7493> which encodes the amino acid sequence <SEQ ID 7494>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2547(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2498

A DNA sequence (GASx491R) was identified in *S.pyogenes* <SEQ ID 7495> which encodes the amino acid sequence <SEQ ID 7496>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-10.24   Transmembrane    6 -  22 (   3 -  28)

----- Final Results -----
             bacterial membrane --- Certainty=0.5097(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2499

A DNA sequence (GASx492) was identified in *S.pyogenes* <SEQ ID 7497> which encodes the amino acid sequence <SEQ ID 7498>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2500

A DNA sequence (GASx493) was identified in *S.pyogenes* <SEQ ID 7499> which encodes the amino acid sequence <SEQ ID 7500>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.69   Transmembrane   21 -  37 (  21 -  37)

----- Final Results -----
               bacterial membrane --- Certainty=0.1277(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2501

A DNA sequence (GASx495R) was identified in *S.pyogenes* <SEQ ID 7501> which encodes the amino acid sequence <SEQ ID 7502>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2891(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2502

A DNA sequence (GASx499R) was identified in *S.pyogenes* <SEQ ID 7503> which encodes the amino acid sequence <SEQ ID 7504>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have an uncleavable N-term signal seq
```

```
INTEGRAL    Likelihood = -2.50   Transmembrane    3 -  19 (   1 -  20)

----- Final Results -----
                bacterial membrane --- Certainty=0.1999(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2503

A DNA sequence (GASx500) was identified in *S.pyogenes* <SEQ ID 7505> which encodes the amino acid sequence <SEQ ID 7506>. Analysis of this protein sequence reveals the following:

```
Possible site: 54

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC77220 GB:AE000497 orf, hypothetical protein [Escherichia
          coli]
 Identities = 262/480 (54%), Positives = 338/480 (69%), Gaps = 5/480 (1%)

Query: 18  GMLNRHGLIAGATGTGKTVTLKVLAEQLSLAGVPVFLADIKGDLSNLTKAGEVTDKLAAR 77
           GM NRHGLI GATGTGKTVTL+ LAE LS  GVPVF+AD+KGDL+ + +AG V++KL AR
Sbjct: 20  GMANRHGLITGATGTGKTVTLQKLAESLSEIGVPVFMADVKGDLTGVAQAGTVSEKLLAR 79

Query: 78  LATIGVSDYQPQAFPVRMWDVFGQNGQPLRTTISELGPMMLSRLLNLNDTQTGVLNIVFK 137
           L  IGV+D+QP A PV +WD+FG+ G P+R T+S+LGP++L+RLLNLND Q+GVLNI+F+
Sbjct: 80  LKNIGVNDWQPHANPVVWDIFGRKGHPVRATVSDLGPLLLARLLNLNDVQSGVLNIIFR 139

Query: 138 IADEKGWLLIDLKDLQAILKEVGDHASDYSSHYGNIAKQSIGAIQRSLLTLEQEGAHQFF 197
           IAD++G LL+D KDL+AI + +GD+A  + + YGNI+  S+GAIQR LL+LEQ+GA  FF
Sbjct: 140 IADDQGLLLLDFKDLRAITQYIGDNAKSFQNQYGNISSASVGAIQRGLLSLEQQGAAHFF 199

Query: 198 GEPALDVADLMQLDVASGVGAINILSATKLPQSPTLVTTFLLWLLSELYKLLPEVGDLDK 257
           GEP LD+ D M+ D A+G G INILSA KL+Q P LY   LLW+LSELY+ LPE GDL+K
Sbjct: 200 GEPMLDIKDWMRTD-ANSKGVINILSARKLYQMPKLVAASLLWMLSELYEQLPEAGDLEK 258

Query: 258 PKMVFFPDEAHLLFKDAPKVFLEKVEQIVRLIRSKGVGIFFVTQNPLDLPETVLAQLGNR 317
           PK+VFFPDEAHLLF DAP+V L+K+EQ++RLIRSKGVG++FV+QNP D+P+ VL QLGNR
Sbjct: 259 PKLVFFPDEAHLLFNDAPQVLLDKIEQVIRLIRSKGVGVWFVSQNPSDIPDNVLQQLGNR 318

Query: 318 IQHAFRAYTPKEQKAVRVAADTFRQNPDLDVARVITELSVGEALISVLNDKGQPSIVERA 377
           +QHA RA+TPK+QKAV+ AA T R NP  D  + I EL  GEALIS L+ KG PS+VERA
Sbjct: 319 VQHALRAFTPKDQKAVKAAAQTMRANPAFDTEKAIQELGTGEALISFLDAKGSPSVVERA 378

Query: 378 YIMPPKSSFAVLSRIESQQLVQSSPFASKYSQSIDRESAYEKLAAEVLEDNNRLAQEAIAT 437
           ++ P S   ++E E   L+  SP   KY   +DRESAYE L  K  + +   Q
Sbjct: 379 MVIAPCSRMGPVTEDENGLINHSPVYGKYDEVDRESAYEML-QKGPQASTEQQNNPPA 437

Query: 438 AQREKEAKEAIKQAATKKANRRSVGRSHKTVVEKATDAFISTTVRTIGRELVRGLLGSL 497
```

```
            +E    + I                    K   +        +  R +  ++VRG+LGSL
Sbjct: 438 KGKEVAVDDGILGGLKDILPGTTGPRGGKK---DGVVQTMAKSAARQVTNQIVRGMLGSL 494
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2504

A DNA sequence (GASx502) was identified in *S.pyogenes* <SEQ ID 7507> which encodes the amino acid sequence <SEQ ID 7508>. Analysis of this protein sequence reveals the following:

```
Possible site: 49

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-13.59   Transmembrane   59 -  75 (  52 -  77)
   INTEGRAL    Likelihood = -9.34   Transmembrane    4 -  20 (   1 -  24)

----- Final Results -----
               bacterial membrane --- Certainty=0.6434(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB15368 GB:Z99121 yvaL [Bacillus subtilis]
 Identities = 28/72 (38%), Positives = 44/72 (60%), Gaps = 2/72 (2%)

Query: 1  MYNLLLTLLVLSGLLSTAIFMQFQKNPSSNVFDSSGSEALFERTKARGFEAFMQRFTAV 60
          M+ +L+T+L+++S  L I + +Q  K+   +   S G+E LF + KARG + + R T V
Sbjct: 1  MHAVLITLLVIVSIALIIVVLLQSSKSAGLSGAISGGAEQLFGKQKARGLDLILHRITVV 60

Query: 61 L--VFPWLAIAL 70
           L  +FF L IAL
Sbjct: 61 LAVLFFVLTIAL 72
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2505

A DNA sequence (GASx505) was identified in *S.pyogenes* <SEQ ID 7509> which encodes the amino acid sequence <SEQ ID 7510>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.44   Transmembrane  140 - 156 ( 138 - 156)

----- Final Results -----
               bacterial membrane --- Certainty=0.1574(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF09704 GB:AE001874 glutamine cyclotransferase [Deinococcus radiodurans]
 Identities = 81/229 (35%), Positives = 128/229 (55%), Gaps = 10/229 (4%)

Query: 16  YSYDSNLYTQGLEQLNNNHILLSAGRYGPSKVGVVDL--TQEIFSEKIAFP-DTVFAEGL 72
```

```
              Y +D    +TQGL+ L    H L S G+ G S + V +L    + ++S   +A       F EG
Sbjct: 54  YPHDRAAFTQGLQYLGGGHYLESTGQVGRSDLRVSELRGAKVLWSTPLAQALPQAFGEGS 113

Query: 73  TTVEDYPWLLTYKEGVAYKFDKATCNCLGAYPFEGDGWGLAYDKENQCLWMTSGNAFLQK 132
           T +     + LT+++GVA   +D  T     G +  ++G+GWGL   D ++   L M++G + L
Sbjct: 114 TQLGSTVYQLTWQDGVALTYDARTFKETGRHRYQGEGWGLTSDGKS--LIMSNGTSTLVW 171

Query: 133 RDPKDFALLDTVLVAIESVPISMLNELEYVDGYLYANIWQTNTIVKLQPDSGKVVATYDI 192
              RDPK FA    +V V  +  P+  LNELEYV G +YAN+W T+ I ++ P +GKV+    D+
Sbjct: 172 RDPKTFAAQRSVQVTDQGQPVRNLNELEYVQSVYANVWLTDRIARIHPQTGKVLTWIDV 231

Query: 193 SPLLKALNLDKSHYPDL----NVLNGIAHLDQQ-RFLITGKLYPLMLEV 236
            S L + ++    +          +V NGIA + ++   L+TGK +P + EV
Sbjct: 232 SDLTREVSAAATKQGQALTFDDVPNGIAFTPSRGYLLLTGKWPTLPEV 280
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2506

A DNA sequence (GASx506R) was identified in *S.pyogenes* <SEQ ID 7511> which encodes the amino acid sequence <SEQ ID 7512>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2800(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2507

A DNA sequence (GASx507R) was identified in *S.pyogenes* <SEQ ID 7513> which encodes the amino acid sequence <SEQ ID 7514>. Analysis of this protein sequence reveals the following:

```
Possible site: 53

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-10.51   Transmembrane  103 - 119 (  97 - 124)
   INTEGRAL    Likelihood = -9.13   Transmembrane  126 - 142 ( 122 - 145)
   INTEGRAL    Likelihood = -8.65   Transmembrane  290 - 306 ( 286 - 307)
   INTEGRAL    Likelihood = -7.17   Transmembrane  200 - 216 ( 198 - 228)
   INTEGRAL    Likelihood = -7.06   Transmembrane   58 -  74 (  54 -  82)
   INTEGRAL    Likelihood = -3.19   Transmembrane  223 - 239 ( 220 - 242)
   INTEGRAL    Likelihood = -2.81   Transmembrane  244 - 260 ( 244 - 261)
   INTEGRAL    Likelihood = -2.71   Transmembrane  174 - 190 ( 169 - 191)

----- Final Results -----
              bacterial membrane --- Certainty=0.5203(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB56669 GB:AL121596 putative membrane protein [Streptomyces
          coelicolor A3(2)]
 Identities = 119/322 (36%), Positives = 182/322 (55%), Gaps = 24/322 (7%)

Query: 9   LETIYILIGLQLSHTAYCFFKDKTNPVYFGTALFWGLLGVTFV----------GGAFL 56
           +E +Y LIGL     A      D++NP  + +A FWGLLG+TF              GG  L
Sbjct: 4   VEWLYWLIGLVFVVMAVQMAMDRSNPKRWTSAAFWGLLGLTFPYGIGVANATAGNGKWTL 63

Query: 57  PNKVIGFIVIVLALLTLFKQVRIGTLPAPNEQKAEESAHRIGNWIPLPVMLMAMISLLLA 116
           P + +G  V+ L +L  F  +  G        ++ E +A R+GN IF+P + + +++++ A
Sbjct: 64  PAEPLGVAVLALIVLAGFNFLGKGVPVTTTGEQREAAAARLGNKIFVPALTIPLVAIVCA 123

Query: 117 LILPDFSKSAIGIAGILA---------TIAILIITKQKPSALLARNNRMNQQVSTSGILP 167
            +L +     G A +L            + +L+  ++K S  +       M + + ++ +LP
Sbjct: 124 SVLDESGLFFTGKATLLGLGLGCVAALVVGMLVTGKKKLSVPIHSGRSMLEAMCSALLLP 183

Query: 168 QLLGALGAIFAAAGVGDVIASLIREIVPADSRFPGVLAVVLGWVIFTMIMGNAFAAFTVI 227
           QLL  LG+IFAAAGVGD +  ++ +++P DS++F VLAY +GM +FT+IMGNAFAAF V+
Sbjct: 184 QLLAVLGSIFAAAGVGDQVGDIMNKVLPDDSKYPAVLAYCVGMFLFTVIMGNAFAAFPVM 243

Query: 228 TTGIGVPFVFAL--GADPIIAGALAMTAGPCGTLLTPMAANFNALPVALMEIKDRNAVIK 285
           T  IG P +     G +P +  A+ M AGF GTL TPMAANFN +P  L+E+KD+   IK
Sbjct: 244 TAAIGWPVLIQQMHGNEPAVL-AIGMLAGFAGTLCTPMAANFNIVPATLLELKDQYGPIK 302

Query: 286 KQAPIALVLIISHIALMYLLAY 307
            Q P  + L+     +M L A+
Sbjct: 303 AQLPTGIALLGCCIVIMALFAF 324
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2508

A DNA sequence (GASx508R) was identified in *S.pyogenes* <SEQ ID 7515> which encodes the amino acid sequence <SEQ ID 7516>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-12.15   Transmembrane  212 - 228 ( 208 - 235)
   INTEGRAL    Likelihood = -8.81   Transmembrane   23 -  39 (  17 -  64)
   INTEGRAL    Likelihood = -7.43   Transmembrane   45 -  61 (  40 -  64)
   INTEGRAL    Likelihood = -1.49   Transmembrane  114 - 130 ( 114 - 130)
   INTEGRAL    Likelihood = -1.49   Transmembrane    3 -  19 (   3 -  20)
   INTEGRAL    Likelihood = -1.49   Transmembrane   76 -  92 (  76 -  92)

----- Final Results -----
               bacterial membrane --- Certainty=0.5861(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB56670 GB:AL121596 possible integral membrane protein
          [Streptomyces coelicolor A3(2)]
 Identities = 77/220 (35%), Positives = 138/220 (62%), Gaps = 2/220 (0%)

Query: 23  IKLIGIVIIVLGFILKCDALATVVVAGLVTALVSGISFIDFLDILGKKFTNQRLLTIFFI 82
           I L+G+V++++LGF+  + + +  V VAG+VT L+   ++ ++ L   G+ F + R +T++ I
Sbjct: 2   IVLLGVVVVILGFVTRRNPVLVVGVAGIVTGLLGKMNPLEVLAAPGRSFADSRSVTVYAI 61

Query: 83  TLPLIGLSETYGLKHRATQLIQRVQALTVCRLLLTLYLIIRELAGLFSIR-LGGHPQFVRP 141
            LP+IGL E YGL+ +A  LI R+  L+ GR LT+YL++R++   F +  +GG  Q VRP
Sbjct: 62  VLPVIGLLERYGLRFQARHLIGRLGKLSAGRFLTVYLLVRQVIAAPGLNSIGGPAQTVRP 121
```

```
Query: 142 LIQPMGEAAAKANIGRELTDAEKDDIKAMAAANENPGNFFAQNTFVGAGGVLLIAGTLEQ 201
           L+ PM EAAA+ + G +L D  ++ +++ +A+ +  G FF ++ F+  G +LLI G +
Sbjct: 122 LVAPMAEAAAAERSTGAKLPDKLREKVRSYGAGADTVGVFFGEDCPIAIGSILLITGFVNS 181

Query: 202 LGY-DGNQAKIAFSSILIAIISIIIVAIYNYLFEKKMERQ 240
               + D     ++A  +I +A+ + +I       L +K++ER+
Sbjct: 182 TYHQDIEPTQLALWAIPLAVCAFLIHGARLLIMDKQQLERR 221
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2509

A DNA sequence (GASx520) was identified in *S.pyogenes* <SEQ ID 7517> which encodes the amino acid sequence <SEQ ID 7518>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2652(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2510

A DNA sequence (GASx522R) was identified in *S.pyogenes* <SEQ ID 7519> which encodes the amino acid sequence <SEQ ID 7520>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2511

A DNA sequence (GASx523) was identified in *S.pyogenes* <SEQ ID 7521> which encodes the amino acid sequence <SEQ ID 7522>. Analysis of this protein sequence reveals the following:

```
Possible site: 22
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2133(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2512

A DNA sequence (GASx525) was identified in *S.pyogenes* <SEQ ID 7523> which encodes the amino acid sequence <SEQ ID 7524>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2364(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2513

A DNA sequence (GASx535) was identified in *S.pyogenes* <SEQ ID 7525> which encodes the amino acid sequence <SEQ ID 7526>. Analysis of this protein sequence reveals the following:

```
Possible site: 47

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4223(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2514

A DNA sequence (GASx536) was identified in *S.pyogenes* <SEQ ID 7527> which encodes the amino acid sequence <SEQ ID 7528>. Analysis of this protein sequence reveals the following:

```
Possible site: 59

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.1102(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB85515 GB:AE000874 conserved protein [Methanobacterium
            thermoautotrophicum]
 Identities = 82/236 (34%), Positives = 132/236 (55%), Gaps = 11/236 (4%)

Query: 9   MNLSIFGLKNIPYLKEGDSIEKLIEESIKTSEFFIEDNDVLCIASKVVSIAEGQVMSLNE 68
           M +S+ G++ +P +  GD I  LI  ++       + D D++ IA  +VS AEG ++SL E
Sbjct: 1   MGISLIGVEGMPLVGAGDDIAYLIISALNEGGEDLLDGDIIVIAETIVSKAEGNIISLEE 60

Query: 69  IQVSDVAKEIHRNIPRKDPRIIEIMNLVNRDLSRLDIKKNYIGCRLENGLKLTSGGIDR 128
           I+ S  A +I      KDP ++E +L         + +  ++I     +G     + GID
Sbjct: 61  IKPSPEALDIAERTG-KDPSLVEAILG---ESSEIIRVGHDFIVSETRHGFVCANAGIDE 116

Query: 129 KSVDEVFL--LPNNPDASAKRISEYLKKSLGKNVAVVITDSDGREDKRGATQVAIGIYGI 186
            +VD+     LP +PD SA++I   L+++ G+ +AV+I+D+ GR  + GA  VA+G+ G+
Sbjct: 117 SNVDDGLATPLPRDPDGSAEKILKTLQEATGRELAVIISDTQGRPRGAVGVAGVAGL 176

Query: 187 HPL--RKTEVIDSQGETIKPQEETLCDMIAACAGLVMGQRGTGIPAVLIRGLDYKW 240
            P+  RK E  D  G +++      + D +AA A LVMGQ   G+PAV+IRG   Y W
Sbjct: 177 SPIWDRKGE-RDLYGRSLETTRVAVADELAAAASLVMGQADEGVPAVIIRG--YPW 229
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2515

A DNA sequence (GASx537) was identified in *S.pyogenes* <SEQ ID 7529> which encodes the amino acid sequence <SEQ ID 7530>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.12   Transmembrane  174 - 190 ( 174 - 190)

----- Final Results -----
                bacterial membrane --- Certainty=0.1447(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2516

A DNA sequence (GASx538) was identified in *S.pyogenes* <SEQ ID 7531> which encodes the amino acid sequence <SEQ ID 7532>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3852(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB99212 GB:U67562 conserved hypothetical protein [Methanococcus
          jannaschii]
 Identities = 129/387 (33%), Positives = 208/387 (53%), Gaps = 44/387 (11%)

Query: 18  EVVERKGLGHPDTLADGIAEQIRIDYSLYCLDKFGVIPHHNFDKLIIRGGHSVQDFGSSD 77
           E+VERKGLGHPD++ DGIAE +          ++KFGI H HN D++ + GGH+   FGG
Sbjct: 20  EIVERKGLGHPDSICDGIAESVSRALCKMYMEKFGTILHHNTDQVELVGGHAYPKPGGGV 79

Query: 78  FIEPIKIIFLGRASKKCFNS------SIPLFKIQKKAATKYLNRILFNLDVENYVEPETL 131
            + PI I+  GRA+ +  +          +P+      KAA +YL ++L N+DV+  V  +
Sbjct: 80  MVSPIYILLSGRATMEILDKEKNEVIKLPVGTTAVKAAKEYLKKVLRNVDVDKDVIID-- 137

Query: 132 TSDPTTKTNWFSPEAIEDLP-EYLDVPKANDTATMISYWPLTISEELALMIEGYFYKLD- 189
                +    S + ++     +  +VP ANDT+  + Y PL+ +E L L  E +    +
Sbjct: 138 -----CRIGQGSMDLVDVFERQKNEVPLANDTSPGVGYAPLSTTERLVLETERFLNSDRL 192

Query: 190 KNELPTPRFTKMGGDIKVMVVRNDLEYSIRINPFLISKFFNNDIESQLYVDKHVEKIKKY 249
           KNE+P     +G DIKVM +R    + ++ I     ++  +  N IE     V    +EK++K
Sbjct: 193 KNEIPA-----VGEDIKVMGLREGKKITLTIAMAVVDRYVHN-IEEYKEV---IEKVRKK 243

Query: 250 IEQKYKNIS--PSIDYH----------YYLTTTGSCIDFGEEGAVGRGNKTHGIISSFR 296
           +E   K I+  + ++ H            YLT TG+  + G++G+VGRGN+ +G+I+ FR
Sbjct: 244 VEDLAKKIADGYEVSIHINTADDYERESVYLTVTGTSAEMGDDGSVGRGNRVNGLITPFR 303

Query: 297 PNTMEAPAGKNCTYFVGKVWGFLSDTIAKEIYEAFNT-PCQIIMQLNIGSKLVRPTHLPI 355
           P +MEA +GKN     VGK++  L++ IA +I +      C + +   IG  +  P  L I
Sbjct: 304 PMSMEAAAGKNPVNHVGKIYNILANLIANDIAKLEGVKECYVRILSQIGKPINEPKALDI 363

Query: 356 Q--TEESVD----QERVLEIVNRHLNN 376
           +  TE+S D    + +  EI N+ L+N
Sbjct: 364 EITDESYDIKDIEPKAKEIANKWLDN 390
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2517

A DNA sequence (GASx539) was identified in *S.pyogenes* <SEQ ID 7533> which encodes the amino acid sequence <SEQ ID 7534>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1436(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2518

A DNA sequence (GASx540) was identified in *S.pyogenes* <SEQ ID 7535> which encodes the amino acid sequence <SEQ ID 7536>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3956(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAD36304 GB:AK001779 conserved hypothetical protein [Thermotoga maritima]
 Identities = 105/353 (29%), Positives = 173/353 (48%), Gaps = 46/353 (13%)

Query: 3   VIGIPTLNEADNISRLVKQIDEYAVNL-GKEIIIINSDSKSTDGTPQIFLETKTYNT-KV 60
           V+GIP+ N A+ IS + +   + V+    + +I+NSD  S DGT + F+ET T+     K
Sbjct: 106 VVGIPSYNNAETISHVARTAAQGIVDFFDDGDGMIVNSDGSGADGTRERFMETDTFGLPKS 165

Query: 61  SIVSEA-KGKGYNVRNIFEYAINHVPNFSGLILIDGDVVSMKKMWLEKMFIAIBSGN-DL 118
           S V E   GKG  +R I E+A+     +   ++ +D D+ S+X  W+E++    +  G  D
Sbjct: 166 SFVYEGLPGKGSAMRAIMEFALKQ--DAEAVVFLDADLRSVKPWWVERLAGPVLKGEADY 223

Query: 119 IIPNVARKSFEGNATNHFIYPMLVKIFKRDMPYQCISGDFGFSRGLIKDLTLKCN--WHK 176
           + P Y R  F+G  TN+  +PM   ++ + +  Q I GDFG  R L++      K    W+
Sbjct: 224 VTPFYLRHRFDGTITNNVCFPMTAVLYGKKVR-QPIGGDFGVGRKLLEEIYLGKPKEIWNT 282

Query: 177 YTLGYGIDIFLTLTAILKGYKIKEIDLQSKIH--KKSFEKIEKLFLEVSQSFTETINDNS 234
              +GIDI++T TAI +S ++ +  L +X+H  K   +  ++FL+V  + FE +
Sbjct: 283 DVARPGIDIWMTTTAINESGRVVQAALGTXVHDVKDPGKHLKGMFLQVVGTLFELV---- 338

Query: 235 LNQDKLRLNINFRESHSRQFIKSSDI----------LSSNDIENLKLRALFLLQEEKQY 282
                I +E+  ++  K  D+              S DI NLK  A   L+E +
Sbjct: 339 ---------ITYENVWKEIWKIEDVPIYGETPQEEVPSMSIDIGNLKKLARETLEEVEYI 389

Query: 283 LHG-LSEVEWDGI--LSNTINNIYRYSSEEHSL-------YLLPLYLLEVYNY 325
            G LSEV+ G   LS+ ++ +YR + +             LLP Y  R   +
Sbjct: 390 DRGILSEVKESGTLSLSSWVDTLYRSAVQYRTRDKKVVENLLPFYFARTARF 442
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2519

A DNA sequence (GASx542) was identified in *S.pyogenes* <SEQ ID 7537> which encodes the amino acid sequence <SEQ ID 7538>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -5.31   Transmembrane    3 -  19 (   1 -  21)

----- Final Results -----
            bacterial membrane --- Certainty=0.3123(Affirmative) < succ>
             bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB07427 GB:AP001519 nucleotide sugar dehydrogenase [Bacillus halodurans]
 Identities = 184/388 (47%), Positives = 274/388 (70%), Gaps = 3/388 (0%)

Query: 1   MKITVVGIGYVGLSIGLLLAKEHDVTFFDIDNKKIDLINKRQSPLKEAAINKLLC-KAKN 59
           M IT+ G GVVGLS  +LLA+ +DV  +DI  +K+D+IN R+SP+ +  I + L  K  N
Sbjct: 1   MNITTAGTGVVGLSNAVLLAQHNDVIAYDIVQEKVDMINNRKSPIVDREIEEFLATKELN 60

Query: 60  INATSSEELAYKDATFIILSLPTNL--KFNKLDTSIIEISVSNILKINKKATIVIKSTVP 117
           + AT+ +E A+KDA F+++S PIN   + N  DTS +E  +S++L IN  A +VIKST+P
Sbjct: 61  LTATTDKEKAFKDAQFVVISTPTNYDPEKNYFDTSSVEAVISDVLSINPNAVMVIKSTIP 120

Query: 118 IGFTEYLRNRFHYNDIIFSPEFLREGSTIHDQLYPSRTIVGNESRNSQLFLDILTDISVE 177
           +G+T  +  RP+  +IIPSPEFLREGS ++D L+PSR +VG  ++ +++F  +L  +++
Sbjct: 121 VGTTREVNERFNTKNIIFSPSFLREGSALYDNLHPSRIVVGERTQRAKIFAALLVQCAIK 180

Query: 178 KDSPSLLVGSSEAEAIKLFSNAYLAQKIAFFNELDTFAEMQNLDSKKIIEAMGYDQRIGN 237
           ++   L   S+EAEAIKLF+N YLA ++AFFNELD++AE++ LD+K+II+ +G D RIG
Sbjct: 181 ENIDVLFTDSTEAEAIKLFANTYLAMRVAFFNELDSYAELKGLDAKQIIDGVGLDPRIGT 240

Query: 238 SHNNPSPGPGGYCLPKDIKQLEYHPKEIPAPIITSISESNLLRKIHIAKMINSSSAKTIG 297
           +NNPSPG+GGYCLPKD KQL  +F+++P  II +I ++N  RK H+A MIL    K +G
Sbjct: 241 HYNNPSPGYGGYCLPKDTKQLLANFEDVPNNIIGAIVDANDTRKDHVANMILKREPKVVG 300

Query: 298 IYRINSKKDSDNCRESSTIDVAKLLKSSGKDVIIFEPLINQKKFLGCPLSNDFNEFIKYS 357
           IYR+  K  SDN R+S+ +DV   L ++G +V++EP ++  +F G  +  DF EF K S
Sbjct: 301 IYRLVMKTGSDNFRQSAILDVMTRLNNAGAEVVVYEPALDATEFDGSKVIEDFAEFKKMS 360

Query: 358 DIIVANRIDDALRKCNSKVFTRDIFQYD 385
           D+IVANR+ D L++   KV+TRD++  D
Sbjct: 361 DVIVANRLSDDLKEVAEKVYTRDLYTRD 388
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2520

A DNA sequence (GASx544R) was identified in *S.pyogenes* <SEQ ID 7539> which encodes the amino acid sequence <SEQ ID 7540>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.06   Transmembrane   46 -  62 (  46 -  62)

----- Final Results -----
                bacterial membrane --- Certainty=0.1022(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2521

A DNA sequence (GASx545R) was identified in *S.pyogenes* <SEQ ID 7541> which encodes the amino acid sequence <SEQ ID 7542>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.49   Transmembrane  186 - 202 ( 186 - 203)

----- Final Results -----
                bacterial membrane --- Certainty=0.1595(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2522

A DNA sequence (GASx546R) was identified in *S.pyogenes* <SEQ ID 7543> which encodes the amino acid sequence <SEQ ID 7544>. Analysis of this protein sequence reveals the following:

```
Possible site: 47

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2422(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2523

A DNA sequence (GASx547R) was identified in *S.pyogenes* <SEQ ID 7545> which encodes the amino acid sequence <SEQ ID 7546>. Analysis of this protein sequence reveals the following:

```
Possible site: 60

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.1612(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2524

A DNA sequence (GASx548) was identified in *S.pyogenes* <SEQ ID 7547> which encodes the amino acid sequence <SEQ ID 7548>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.5156(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2525

A DNA sequence (GASx549R) was identified in *S.pyogenes* <SEQ ID 7549> which encodes the amino acid sequence <SEQ ID 7550>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
              bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2526

A DNA sequence (GASx552) was identified in *S.pyogenes* <SEQ ID 7551> which encodes the amino acid sequence <SEQ ID 7552>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.59   Transmembrane   83 -  99 (  83 -  99)

----- Final Results -----
             bacterial membrane --- Certainty=0.1235(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2527

A DNA sequence (GASx553) was identified in *S.pyogenes* <SEQ ID 7553> which encodes the amino acid sequence <SEQ ID 7554>. Analysis of this protein sequence reveals the following:

```
Possible site: 49

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2781(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2528

A DNA sequence (GASx554) was identified in *S.pyogenes* <SEQ ID 7555> which encodes the amino acid sequence <SEQ ID 7556>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2792(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2529

A DNA sequence (GASx555) was identified in *S.pyogenes* <SEQ ID 7557> which encodes the amino acid sequence <SEQ ID 7558>. Analysis of this protein sequence reveals the following:

```
Possible site: 35

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.00   Transmembrane   49 -  65 (  49 -  65)

----- Final Results -----
              bacterial membrane --- Certainty=0.1001(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA36631 GB:AB016282 ORF25 [bacteriophage phi-105]
 Identities = 43/118 (36%), Positives = 69/118 (58%), Gaps = 2/118 (1%)

Query: 3   LLDLIGRKRARDKPQNSYEGQDFSYLPG--RTTSGRNVDEFKIMQTTAVYACVRVLAEAV 60
            LL+ +  KR+              +FG  +T SGE V E  ++     ++ACV VL++ +
Sbjct: 2   LLERMPEKRSGSSSDHEDGFNNILLNMPGGRKTASGKRVSESNSLVQPDIFACVNVLSDDI 61

Query: 61  ASLPIHIYERTENGKEKKLDHPLYFLLHDEPNPEMSSPIPRETIMSHLLIWGNAYVQI 118
            A LPIH Y+RT+ G E+K +H      ++  PNP M++P +++ +M+H+L WGNAY  I
Sbjct: 62  AKLPIHTYKRTDGGIERKPEHKSAHAVYARPNPYMTAFTWKKIMMTHVLTWGNAYSYI 119
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2530

A DNA sequence (GASx556) was identified in *S.pyogenes* <SEQ ID 7559> which encodes the amino acid sequence <SEQ ID 7560>. Analysis of this protein sequence reveals the following:

```
Possible site: 43

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2055(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2531

A DNA sequence (GASx557) was identified in *S.pyogenes* <SEQ ID 7561> which encodes the amino acid sequence <SEQ ID 7562>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1696(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2532

A DNA sequence (GASx559) was identified in *S.pyogenes* <SEQ ID 7563> which encodes the amino acid sequence <SEQ ID 7564>. Analysis of this protein sequence reveals the following:

```
Possible site: 51

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1556(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB15798 GB:Z99123 alternate gene name: ipa-83d [Bacillus subtilis]
 Identities = 70/263 (26%), Positives = 121/263 (45%), Gaps = 25/263 (9%)

Query: 68  KTIEQIKELK--YSIDAVACWDEALTHIADDISKELGLNPISSLDSQSFRFKDRMRMVCE 125
          + +EQI ++   +  DA+   +E          + LGL       +++ R K++MR
Sbjct: 87  EVVEQIVKVAEMPGADATITNNELFIAPMAKACERLGLKGAGVQAAENARDKNKMRDAFN 146

Query: 126 AGGLKMPKVKIINQFSDTNKIINW-KYPLIVKPTSFLASIGVKKVYNPSELQQAVSQMLN 184
            G+K  K K +    D   +    PLI+KPT   +SIGV  + +    +  +++ +
Sbjct: 147 KAGVKSIKNKRVTTLEDPRAALEEIGTPLILKPTYLASSIGVTLITDTETAEDEFNRVND 206

Query: 185 VKPPVYIASGVYELGELVNLEPRVLVEEFIDGE----------EY-SLESVVRNGIYTP 232
              + +   V         E   + EEF+ GE           +Y S+E ++ +G Y P
Sbjct: 207 YLKSINVPKAV-------TFEAPPIAEEFLQGEYGDMYQTEGYSDYISIEGIMADGEYFP 259

Query: 233 LGITKKIVDEKLFMDEIGHIPPSNLNKEEKSRVYSWAEKLHQILQLNHITTHTEFRIGRN 292
          + I  K    ++   E  HI PS L++E K ++   A+K ++ L L +  THTE ++ +N
Sbjct: 260 YAIHDKT--PQIGPTETSHITPSILDEEAKKKIVEAAKKANEGLGLQNCATHTEIKLMKN 317

Query: 293 GDIILIEIGARIGG-DCIPNLMK 314
            + LIE  AR  G + IPN+ K
Sbjct: 318 REPGLIESAARFAGWNMIPNIKK 340
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2533

A DNA sequence (GASx561) was identified in *S.pyogenes* <SEQ ID 7565> which encodes the amino acid sequence <SEQ ID 7566>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2602(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2534

A DNA sequence (GASx562) was identified in *S.pyogenes* <SEQ ID 7567> which encodes the amino acid sequence <SEQ ID 7568>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
           bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAD06696 GB:AE001539 HISTIDYL-TRNA SYNTHETASE [Helicobacter
          pylori J99]
 Identities = 75/309 (24%), Positives = 129/309 (41%), Gaps = 35/309 (11%)

Query: 11  KGYHRQFNQILLGAWGIESAYVDAEIIVATWRGLQRPKGIKVE--FIQLSNKNIFDVLEK 68
           KG  R+F Q      G ES   DAEII       L   K + +E   + ++++ I + + +
Sbjct: 115 KGRYREFTQCDFDFIGSESSLVCDAEIIQVIIASL---KALDLEDFCVSINHRKILNGIC8 171

Query: 69  DLSKKLRFEDISIEAILGKYLCNNDIEIIKCLYEKDKINMELLISLISKISNKLVKQEPI 128
                E + I   L K   N   E +K   + D   ++ L+ ++     N L   EP
Sbjct: 172 YFGIAQVNEVLRIVDKLEKIGLNGVEEELKKECDLDSNTIKDLLEMVQIKQNDLSHAEFF 231

Query: 129 -KVLVLYEYVKNFLP----VDCIYPSLS------NLY--------GTGHYSSMNYKIFIR 169
             K+  L +Y +N         ++ +Y  L        NLY         G G+Y+ + Y+  +
Sbjct: 232 EKIAYLKDYNENLKKGIQDLERLYQLIGDLQISQNLYKIDFSIARGLGYYTGIVYETTLN 291

Query: 170 TKSGDIFDIADGGRIDDMVSKFNKVNVLGVCMGIGTTVLSQEI-------EYEIEDRIMI 222
             +  + GGR D +   F+X N+ GV   IG   L   +        E   + +++I
Sbjct: 292 DMKS-LGSVCSGGRYDHLTKNFSKENLQGVGASIGIDRLIVALSEMQLLDERSTQAKVLI 350

Query: 223 LVEKIDVKIYKNCLELANKLSGYHCSVFEFPYKKIKKFPKHELYSRHHYIIVRLDGSMEY 282
             +   Y N L  + + SG    V+    +KIKK F +  +  H ++ V   G  E+
Sbjct: 351 ACMHEEYFSYANRLAESLRQSGIFSEVYP-EAQKIKKPFSYANHKGHEFVAV--IGEEEF 407

Query: 283 RFSSVALKN 291
           +  +++LKN
Sbjct: 408 KSETLSLKN 416
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2535

A DNA sequence (GASx564) was identified in *S.pyogenes* <SEQ ID 7569> which encodes the amino acid sequence <SEQ ID 7570>. Analysis of this protein sequence reveals the following:

```
Possible site: 56

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1264(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2536

A DNA sequence (GASx576) was identified in *S.pyogenes* <SEQ ID 7571> which encodes the amino acid sequence <SEQ ID 7572>. Analysis of this protein sequence reveals the following:

```
Possible site: 60

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2537

A DNA sequence (GASx577R) was identified in *S.pyogenes* <SEQ ID 7573> which encodes the amino acid sequence <SEQ ID 7574>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence
   INTEGRAL     Likelihood = -2.60   Transmembrane     2 -  18 (   1 -  18)

----- Final Results -----
              bacterial membrane --- Certainty=0.2041(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2538

A DNA sequence (GASx579) was identified in *S.pyogenes* <SEQ ID 7575> which encodes the amino acid sequence <SEQ ID 7576>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial cytoplasm --- Certainty=0.3161(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB12286 GB:Z99106 similar to hypothetical proteins [Bacillus subtilis]
 Identities = 62/140 (44%), Positives = 88/140 (62%), Gaps = 3/140 (2%)

Query: 3   LYTNYVQEVSLADPGKPLHHKAYWNKRLKTTGGRFFPKDGHLDFNPRMLKEHGELIFRKIV 62
            L    +++S   F KP  H+A +N RLKTTGGR+      +++ N + L EHG      I+
Sbjct: 6   LQKLTEDISETYFKKPPRHQALFNDRLKTTGGRYLITSHNIELNRKYLIEHGREELIGII 65

Query: 63  RHELCHYHLYFEGRGYHHKDRDFKDLLAQVNGLRY---VPTSSKSKTNHHYSCQTCGQVY 119
            +HELCHYHL+ EG+GY H+DRDF+ LL QVN  R+   +   +++K  + Y C TCGQ Y
Sbjct: 66  KHELCHYHLHLEGKGYKHRDRDFRMLLQQVNAPRFCTPLKKKAENKKTYMYICTTCGQQY 125

Query: 120 QRKRRINLAKYVCSNCHGKL 139
             +KR +N  +Y C  C GK+
Sbjct: 126 IKKRAMNFDRYRCGKCRGKI 145
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2539

A DNA sequence (GASx587R) was identified in *S.pyogenes* <SEQ ID 7577> which encodes the amino acid sequence <SEQ ID 7578>. Analysis of this protein sequence reveals the following:

```
Possible site: 53

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-10.40   Transmembrane    46 -  62 (  39 -  89)
   INTEGRAL    Likelihood = -5.36   Transmembrane    65 -  81 (  63 -  89)

----- Final Results -----
                bacterial membrane --- Certainty=0.5161(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2540

A DNA sequence (GASx590R) was identified in *S.pyogenes* <SEQ ID 7579> which encodes the amino acid sequence <SEQ ID 7580>. Analysis of this protein sequence reveals the following:

```
Possible site: 35

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2036(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2541

A DNA sequence (GASx592R) was identified in *S.pyogenes* <SEQ ID 7581> which encodes the amino acid sequence <SEQ ID 7582>. Analysis of this protein sequence reveals the following:

```
Possible site: 23

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -4.62   Transmembrane   25 -  41 (  24 -  43)

----- Final Results -----
                bacterial membrane --- Certainty=0.2848(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2542

A DNA sequence (GASx600) was identified in *S.pyogenes* <SEQ ID 7583> which encodes the amino acid sequence <SEQ ID 7584>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -2.18   Transmembrane    3 -  19 (   2 -  19)

----- Final Results -----
                bacterial membrane --- Certainty=0.1871(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2543

A DNA sequence (GASx603R) was identified in *S.pyogenes* <SEQ ID 7585> which encodes the amino acid sequence <SEQ ID 7586>. Analysis of this protein sequence reveals the following:

```
Possible site: 48

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.3027(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA03927 GB:AJ000109 gluthatione peroxidase [Lactococcus lactis]
 Identities = 79/133 (59%), Positives = 103/133 (77%)

Query: 1   VVLVVNTATKCGLTPQYQALQALVDTYHDKGFEVLDFPCNQFLNQAPGDAEEINHFCSLT 60
           VV+VVNTA+KCG TPQ++ L+ LY+TY D+G E+L FPCNQF NQ  G+  EIN FC L
Sbjct: 25  VVIVVNTASKCGFTPQFEGLEKLYETYKDQGLEILGFPCNQFANQDAGENTEINEFCQLN 84

Query: 61  YHTTFPRFAKIKVNGKDADPLFTWLKEEKSGPLGKRIEWNFTKFLIDQNGQVIKRYSSKT 120
           Y  TF  F KIKVNGK+A PL+ +LK+E  G L   I+WNFTKFLID+GQVI+R++ KT
Sbjct: 85  YGVTFTMFQKIKVNGKEAHPLYQFLKKEAKGALSGTIKWNFTKFLIDRDGQVIERFAPKT 144

Query: 121 DPKLIEEDLKALL 133
           +P+ +EE++K LL
Sbjct: 145 EPEEMEEIKKLL 157
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2544

A DNA sequence (GASx605) was identified in *S.pyogenes* <SEQ ID 7587> which encodes the amino acid sequence <SEQ ID 7588>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3687(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2545

A DNA sequence (GASx608R) was identified in *S.pyogenes* <SEQ ID 7589> which encodes the amino acid sequence <SEQ ID 7590>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1327(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2546

A DNA sequence (GASx616) was identified in *S.pyogenes* <SEQ ID 7591> which encodes the amino acid sequence <SEQ ID 7592>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2547

A DNA sequence (GASx617R) was identified in *S.pyogenes* <SEQ ID 7593> which encodes the amino acid sequence <SEQ ID 7594>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial cytoplasm --- Certainty=0.0677(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2548

A DNA sequence (GASx622R) was identified in *S.pyogenes* <SEQ ID 7595> which encodes the amino acid sequence <SEQ ID 7596>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL     Likelihood = -7.32   Transmembrane     4 -  20 (   1 -  26)

----- Final Results -----
              bacterial membrane --- Certainty=0.3930(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2549

A DNA sequence (GASx632) was identified in *S.pyogenes* <SEQ ID 7597> which encodes the amino acid sequence <SEQ ID 7598>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -3.40   Transmembrane   83 -  99 (  82 - 102)
   INTEGRAL    Likelihood = -1.28   Transmembrane  108 - 124 ( 108 - 124)

----- Final Results -----
                bacterial membrane --- Certainty=0.2359(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2550

A DNA sequence (GASx638) was identified in *S.pyogenes* <SEQ ID 7599> which encodes the amino acid sequence <SEQ ID 7600>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -0.64   Transmembrane   12 -  28 (  12 -  28)

----- Final Results -----
                bacterial membrane --- Certainty=0.1256(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2551

A DNA sequence (GASx652R) was identified in *S.pyogenes* <SEQ ID 7601> which encodes the amino acid sequence <SEQ ID 7602>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2622(Affirmative) < succ>
```

```
bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA74610 GB:Y14232 hypothetical protein [Bacteriophage TP901-1]
 Identities = 225/485 (46%), Positives = 308/485 (63%), Gaps = 20/485 (4%)

Query: 2   RKVAIYSRVSTINQAKEGYSIQGQIKALTKYCKAMEWKIYKNYSDAGFSGGKLKRKPAITE 61
            +KVAIY+RVST NQAKEG+SI  QI+ LTKY EAM W++    Y+DAGFSG KLERPA+
Sbjct: 3   KKVAIYTRVSTTNQAKEGFSIDEQIDRLTKYAEAMGWQVSDTYTDAGFSGAKLERPAMQR 62

Query: 62  LIEDGKNNKFDTILVYKLDRLSRNVKDTLYLVKDVFTANNIHFVSLKENIDTSSAMGNLF 121
            LI D +N  FDT+LVYKLDRLSR+V+DTLYLVKDVFT N I F+SL E+IDTSSAMG+LF
Sbjct: 63  LINDIENKAFDTVLVYKLDRLSRSVKDTLYLVKDVFTKNKIDFISLNESIDTSSAMGSLF 122

Query: 122 LTLLSAIAEFERELIKERMQFGVMNRAKSGKTTAWKTPPYGYRYNKDEKTLSVNELEAAN 181
            LT+LSAI EFERE IKERM  G +  RAKSGK+  W     +GY +N+     L +  L+A
Sbjct: 123 LTILSAINEFERENIKERMTMGKLGRAKSGKSMWWTKTAPGYYHNRKTGILEIVPLQATI 182

Query: 182 VRQMFDMIISGCSIMSITNYARDN-FVGN--TWTHVEVKRILENETTYKGLVKYRBQTPSG 238
            V Q+F   +SG S+  + +   ++  +G    W++  +++ L+N  Y G +K+++  F G
Sbjct: 183 VEQIFTDYLSGISLTKLRDKLNESGHIGKDIPWSYRTLRQTLDNPVYCGYIKFKDSLFBG 242

Query: 239 DHQAIIDEKTYNKAQIALAHRT----DTKTNIRPFQGKYMLSHIAKCGYCGAPALKVCTGR 294
             H+ II  +TY K Q  L  R      +   N RPFQ KYMLS +A+CGYCGAPLK+  G
Sbjct: 243 MHKPIIFYRTYLKVQKELEERQQQTYTERNNNPRPFQAKYMLSGMARCGYCGAPLKIVLGR 302

Query: 295 AKNDGTRRQTYVCVNKTESLARRSVNNYNNQKICNTGRYEKKHIEKYVIDVLVKLQKDKE 354
             + DG+R   Y C N+      + +  YN+ K C++G Y+  ++E  VID L   Q + +
Sbjct: 303 KRKDGSRTMKYHCANRFPR-KTKGITVYNDNKKCDSGTYDLSNLENTVIDNLIGFQENND 361

Query: 355 YLKKIKKDDN--IIDITPLKKEIRIIDKKINKLNDLYINDLIDLPKLKKDIEELNHLKDD 412
             L KI   +N  I+D +  KK+I  IDKKI + +DLY+ND I + +LK   + L   K
Sbjct: 362 SLLKIINGNNQPILDTSSFKKQISQIDKKIQKNSDLYLNDFITMDELKDRTDSLQAEK-- 419

Query: 413 YNKAIKLNYLDKKNEDSLGML------MINLDIRKSSYDVQSRIVKQLIDRVEVTMDNID 466
              K +K   + K  DS  +        + ++ I + SYD + +IV  L+ +V+VT DN+D
Sbjct: 420 --KLLKAKISENKFNDSTDVPELVKTQLGSIPINELSYDNKKKIVNNLVSKVDVTADNVD 477

Query: 467 IIFKF 471
            IIFKF
Sbjct: 478 IIFKF 482
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2552

A DNA sequence (GASx653R) was identified in *S.pyogenes* <SEQ ID 7603> which encodes the amino acid sequence <SEQ ID 7604>. Analysis of this protein sequence reveals the following:

```
Possible site: 48

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.22   Transmembrane   86 - 102 (  86 - 102)

----- Final Results -----
              bacterial membrane --- Certainty=0.1489(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF12707 GB:AF066865 unknown [bacteriophage TPW22]
 Identities = 45/67 (67%), Positives = 53/67 (78%), Gaps = 2/67 (2%)

Query: 57  EKEAVRCPKCKSTNVGPMQQGKKTFSVKKAVAGTLLIG--GTVMGPLGEKGKKQWHCNEC 114
           +K A++CPKCKST+V FMQQGKK FSV KAV G +L G  GT+ GP G+KGKKQWHCN C
Sbjct: 138 DKHAIKCPKCKSTDVVFMQQGKKGPSVGKAVGGAVLTGGIGTLAGPAGKKGKKQWHCNNC 197

Query: 115 SCIFETK 121
           +FETK
Sbjct: 198 GRVFETK 204
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2553

A DNA sequence (GASx655) was identified in *S.pyogenes* <SEQ ID 7605> which encodes the amino acid sequence <SEQ ID 7606>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3956(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB63661 GB:AJ251789 Cro protein [Lactobacillus casei
            bacteriophage A2]
 Identities = 43/76 (56%), Positives = 55/76 (71%)

Query: 26  MTINLKRLKAERIASGMTQCEVAQSMGWKTRTPVAKRENGIVSIGADELAKITLIPGLPI 85
           MT+NLKRL+AERIA GM Q E+A++MGW TR+ VAKRENGI +I A KL K+  I G
Sbjct: 1   MTLNLKRLRASRIAKGMNQDEMAKAMGWHTRSSVAKRENGITTISATELVKMASILGYGT 60

Query: 86  EKIAIFPDKDVPVMER 101
            ++ +FF  +VP  ER
Sbjct: 61  NQLDLFFTNNVPDRER 76
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2554

A DNA sequence (GASx656) was identified in *S.pyogenes* <SEQ ID 7607> which encodes the amino acid sequence <SEQ ID 7608>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4505(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2555

A DNA sequence (GASx657) was identified in *S.pyogenes* <SEQ ID 7609> which encodes the amino acid sequence <SEQ ID 7610>. Analysis of this protein sequence reveals the following:

```
Possible site: 35

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.6593(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2556

A DNA sequence (GASx658) was identified in *S.pyogenes* <SEQ ID 7611> which encodes the amino acid sequence <SEQ ID 7612>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.5244(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2557

A DNA sequence (GASx660) was identified in *S.pyogenes* <SEQ ID 7613> which encodes the amino acid sequence <SEQ ID 7614>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1133(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB99331 GB:U67572 purine NTPase [Methanococcus jannaschii]
 Identities = 71/346 (20%), Positives = 154/346 (43%), Gaps = 52/346 (15%)

Query: 8   MSITTINKLEIENVK-----RIKAVKIEPGATGLYIIGGNNNCGKTSVLDAIAWAL--GGN 60
           MS+ + ++ + N K      RIK  K       G+  I G N  GK+S+ +A+ +AL   G+
Sbjct: 1   MSMILKEIRMNNFKSHVNSRIKPEK------GIVAIIGENGSGKSSIFEAVFFALFGAG 54

Query: 61  KYKPSQAMEGSQ---VPFTLKITMSNGLIVERKGKNASLKVIDPNGQ----------KG 107
            +     + +G +   V    ++  +N  I+            + NG+          K
Sbjct: 55  NFNYDTIITKGKKSVYVELDFEVNGNNYKIIREYDSGRGGAKLYKNGKPYATTISAVNKA 114

Query: 108 GQQLL----DSFVEELAI---NLPKFMDSTPKHADVLLEIIGVGDQLAELKEKEIYN 160
             ++L    + F+  + I    + KF+   P EK + + +++G+ D+  +   K  EI
Sbjct: 115 VNEIIGVDRNMFLNSIYIKQGEIAKFLSLKPSEKLEVVAKLLGI-DEFEKCYQKMGEIVK 173

Query: 161 QRHAIGVIADQKEKFAKEMTYYPDAPKQLVS-ISELIQQHQAILAKNGE-NAQKR--QNV 216
           +          + E+   E+ Y  +  K+L + +S+L ++++ ++  N + N  K+  +++
Sbjct: 174 E------YEKRLSRIEGELNYKENYEKELKNKMSQLEEKNKKLMEINDKLNKIKKEFEDI 227

Query: 217 ERIRYDYNQSILEVDRLRKLLADAEAKTNKLSEDLKIANTD------AMDLHDESTAEIE 270
           E++  ++    L  ++    L + +       +++LKI   D      A +  +   E E
Sbjct: 228 EKLFNEWENKKLLYEKFINKLEERKRALAELKNQELKLEYDLNTVVEARETLNRHKDEYE 287

Query: 271 ANIADIDEVNRKVRANFDKDKAE-EDAKQQREQYNILTNDIESIRQ 315
            + +DE+ RK+ +   + K+  ED  +  +Q  I+  DIE +++
Sbjct: 288 KYKSLVDEI-RKISSRLELKSHYEDYLKLITQLEIIKGDIEKLKE 332
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2558

A DNA sequence (GASx661) was identified in *S.pyogenes* <SEQ ID 7615> which encodes the amino acid sequence <SEQ ID 7616>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1559(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2559

A DNA sequence (GASx662) was identified in *S.pyogenes* <SEQ ID 7617> which encodes the amino acid sequence <SEQ ID 7618>. Analysis of this protein sequence reveals the following:

```
Possible site: 52
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.3292(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2560

A DNA sequence (GASx663) was identified in *S.pyogenes* <SEQ ID 7619> which encodes the amino acid sequence <SEQ ID 7620>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.4867(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2561

A DNA sequence (GASx664) was identified in *S.pyogenes* <SEQ ID 7621> which encodes the amino acid sequence <SEQ ID 7622>. Analysis of this protein sequence reveals the following:

```
Possible site: 46

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2141(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2562

A DNA sequence (GASx667) was identified in *S.pyogenes* <SEQ ID 7623> which encodes the amino acid sequence <SEQ ID 7624>. Analysis of this protein sequence reveals the following:

```
Possible site: 59

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2614(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF80834 GB:AF165214 Orf78 [Pseudomonas phage D3]
 Identities = 68/200 (34%), Positives = 109/200 (54%), Gaps = 10/200 (5%)

Query: 12  GLRFGSLTVINRNRNNSKGGNARWNCLCDCGNKTVVI-GSKLRSGYTKSCGCARKNDNAK 70
           GLR G + V      ++  G  + W C CDCGN+ ++  G+ +R+  T SCGC+R +
Sbjct: 8   GLRVGKVVV--EAFSHCAGKASHWVCRCDCGNRVIMRRGNLMRNRTTTSCGCSRFSH--- 62

Query: 71  GYSSTRLYRIWKGMMNRCYNHKNDNYKYYGGKGISICDEWLTPINFRTWSLSNGYKESLT 130
           G + T  Y  W  M++RC N  N  Y  Y G+GI++C+ W+TF NF        G +   T
Sbjct: 63  GMTGTFTYSSWSNMIDRCTNPSNKRYVDYQGRGITVCERWMTFANFLA---DMGERPDAT 119

Query: 131 -IDRINPKGNYTPLNCRWVSMIMQQNNKTNNRYLSYLGQETTIAEFSEKLNVTYWTVINQ 189
            +DRI+     Y   NCRW +   Q NN   N ++ YLG+  T+++++ +L +    T+ ++
Sbjct: 120 SLDRIDNDAGYFKENCRWATALEQMNNTRRNTFVEYLGRRQTVSQWAGQLGIPECTLRSR 179

Query: 190 LKLGWSVERIVEEARMKNDR 209
           L  GWS+E  +++    K  R
Sbjct: 180 LNRGWSIEDAMQKPISKQRR 199
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2563

A DNA sequence (GASx668) was identified in *S.pyogenes* <SEQ ID 7625> which encodes the amino acid sequence <SEQ ID 7626>. Analysis of this protein sequence reveals the following:

```
Possible site: 41

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1476(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB75598 GB:AJ271879 putative DNA helicase [uncultured
           eubacterium]
 Identities = 42/168 (25%), Positives = 75/168 (44%), Gaps = 7/168 (4%)

Query: 374 IAGPSKAGKSFALIELSIALAEQGKWLG-WQCEQGKVLVYNLELDRPGALHRFKDVYDAM 432
           +  P  AGKS   ++L+  +A G   LG  +   G V+Y+  E D P+A+H       A
Sbjct: 35  LVSPGGAGKSMLALQLAAQIAGGFDLLGVGELPTGPVIYLPAE-DPPTAIHHRLHALGAH 93
```

```
Query: 433 GLPPANVANIDIWNLRGKTVPMDKLAPKLIRRSLKKNYQA---VIIDPIYKVLTGDENSA 489
                 A  D   ++      +  +          +LK+  +     +I+D + +     +EN++
Sbjct: 94  LSAEERQAVADGLLIQPLIGSLPNIMASNWFEALKRAAEGRRLMILDTLRRFHIEEENAS 153

Query: 490 DQMAFFTNQFDKVATELGCSVIYCHHHSXGS--QGGKKSMDRASGSGV 535
              MA     + + +A + GCS+++ HH SXG+   G        + GS V
Sbjct: 154 GPMAQVIGRMEAIAADTGCSIVFLHHASKGATMWGAGDQQQASRGSSV 201
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2564

A DNA sequence (GASx669) was identified in *S.pyogenes* <SEQ ID 7627> which encodes the amino acid sequence <SEQ ID 7628>. Analysis of this protein sequence reveals the following:

```
Possible site: 56

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2555(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2565

A DNA sequence (GASx670) was identified in *S.pyogenes* <SEQ ID 7629> which encodes the amino acid sequence <SEQ ID 7630>. Analysis of this protein sequence reveals the following:

```
Possible site: 54

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2921(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF74082 GB:AF212845 ORF129 [Lactococcus lactis bacteriophage
            ul36]
 Identities = 36/108 (33%), Positives = 63/108 (58%), Gaps = 1/108 (0%)

Query: 8   IEFFLPMDKIPTTTHQQKKVTVINGKPHFYEPSSLKNARDKFTSLLAQHVPPSKLDGPIR 67
            ++F   +DK+PTT  QQK +  + GK  FY+     KN   K   +  +       + + P++
Sbjct: 1   MKFEPELDKMPTT-QQQKGIKKVKGKLQFYDRRGTKNYSLKAQLMKNKPKRCFEKNVPLK 59

Query: 68  LTVKWLFPKIKGSTNGQYKTTKPDTDNLQKLLKDCWTELGPWNDDAQV 115
            L+V + +    +      Q+KT++PD  DNL K L+D MT+L +++DD+Q+
Sbjct: 60  LSVTFFYAIKQKKRWWQWKTSRPDLDNLMKNLQDYMTKLRYYSDDSQI 107
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2566

A DNA sequence (GASx671) was identified in *S.pyogenes* <SEQ ID 7631> which encodes the amino acid sequence <SEQ ID 7632>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4294(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2567

A DNA sequence (GASx672R) was identified in *S.pyogenes* <SEQ ID 7633> which encodes the amino acid sequence <SEQ ID 7634>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -6.37   Transmembrane  106 - 122 ( 104 - 125)

----- Final Results -----
            bacterial membrane --- Certainty=0.3548(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2568

A DNA sequence (GASx673) was identified in *S.pyogenes* <SEQ ID 7635> which encodes the amino acid sequence <SEQ ID 7636>. Analysis of this protein sequence reveals the following:

```
Possible site: 56

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4781(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB18697 GB:U38906 ORF22 [Bacteriophage r1t]
 Identities = 78/207 (37%), Positives = 123/207 (58%), Gaps = 2/207 (0%)

Query: 28  EIHRILGIDEVYKAPKRLTDILPDKDSREDIFRQFLKYETDVSYDWFMQYFEEEQADRKN 87
           + + +L +DE      R+ +++FDK  RE+ + + L    D+  D+F  YF    A
Sbjct: 7   QFYDMLNVDEHMNFTNRIQKLVFDKKGREEFYSKILNIHHDMGVDFFRDYFMAHSAVSA- 65

Query: 88  KKQDFTPKSVSTLLSKIISGNQYYEVA-VGTGGILIQAWQEQRLNDSPFTYRPSKYWYHV 146
           K Q +TP  +  L + ++ G+   ++    GTG ++IQ WQ+ R+N   F Y PS YWY
Sbjct: 66  KGQHYTPDELGKLTALLVGSGSGADLTGAGTGTLIIQKWQDDRMNTDFFNYLPSNYWYQA 125

Query: 147 SELSDKAVPFLLFNMSIRGINGVVVHGDSLTRQVENIYFLQNTKDDMLSKSDINVMPRTQ 206
           ELSD+A+ FL+   +IRG+NGVV+HGD+L   VK +YF+CN+ ++ + FS+INV+P ++
Sbjct: 126 LELSDEAISFLIHAFAIRGMNGVVIHGDALEMAVKQVYFIQNSANNPIGFSEINVIPHSK 185

Query: 207 DIEREPNVKSWIGDGIEHIRNPLIEWI 233
           D     + EW    IEHIE+   +WI
Sbjct: 186 DAMEFLGIHEWTSQAIEHIESKFPDWI 212
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2569

A DNA sequence (GASx674) was identified in *S.pyogenes* <SEQ ID 7637> which encodes the amino acid sequence <SEQ ID 7638>. Analysis of this protein sequence reveals the following:

```
Possible site: 51

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.00   Transmembrane  122 - 138 ( 122 - 138)

----- Final Results -----
               bacterial membrane --- Certainty=0.1001(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF63071 GB:AF158600 gp137 [Streptococcus thermophilus
           bacteriophage Sfi11]
 Identities = 66/135 (48%), Positives = 89/135 (65%), Gaps = 2/135 (1%)

Query: 5   PEIDIQKTKSNAKRKLREYPRWRRIANDVDTQKVTATYSFEPRQSHGVPSKPVERLALNR 64
           PEID + T    KRKLREYPRWR IA+D   QK+T  ++F PR   G  +KPVE +A+ R
Sbjct: 4   PEIDEKATLKRCKRKLREYPRWRELAHDSAEQKITQEFTPMPRG--GGVNKPVENIAVER 61

Query: 65  VSAEQELDAIEQAVSMILEPERRRILYDKYLAPYKKADKVIYTELCMSESFYYDTLDIAL 124
           V A  EL+AIEQAV+ +  P+ RRIL +KYLA   K +  I   +  + + + L+ ++
Sbjct: 62  VDALNELEAIEQAVNGLYRPDYRRILIEKYLAYPFKPNWQIAQSIFERTAFQELINNSI 121

Query: 125 LAFAELYREGVLLVE 139
           LAFAELYR+G L+VE
Sbjct: 122 LAFAELYRDGRLIVE 136
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2570

A DNA sequence (GASx675) was identified in *S.pyogenes* <SEQ ID 7639> which encodes the amino acid sequence <SEQ ID 7640>. Analysis of this protein sequence reveals the following:

```
Possible site: 41

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1865(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2571

A DNA sequence (GASx676) was identified in *S.pyogenes* <SEQ ID 7641> which encodes the amino acid sequence <SEQ ID 7642>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4870(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB07254 GB:AP001519 unknown [Bacillus halodurans]
 Identities = 194/451 (43%), Positives = 262/451 (58%), Gaps = 69/451 (15%)

Query: 1   MEFVDKKLSEITPYQNNPRNNDEAVGPVAE----SIKERPGPKVPIVV-DKWGRIVNGHTR 55
           +  V+KK+ ++ P + NPR + +    P  E    SI+EPG   PIV  ++ G +V GH R
Sbjct: 3   IRIVNKKIDDLVPAEYNPRLDLQPGDPEYEKLKRSIEEPGLVEPIVFNERTGRVVGGHQR 62

Query: 56  YKAAQKLGLETVPVIVADDLSKEQIKAFRLADNKV-GEIAVWDLDLLNEELNDILDLDMS 114
            K  ++LG E VPV V D L +   KA  +A NK+ G+    + L  L EEL+  L +D++
Sbjct: 63  LKILRRLGWKEVPVSVVD-LDDHHEKALNVALNKIEGDWDNFKLKELLEELDSGL-IDVT 120

Query: 115 APGFDVLDNLDDL-----IEDEKDL--DDF----TGTVPDEPKSKLGDIYQLGSHKLMCG 163
             GFD  + ++DL     +EDE ++  DDF           +EP +K GD++ LG H L+ G
Sbjct: 121 LTGFDE-EEIKDLMTQFFVEDENEIKEDDFDPEVAEEIEEPITKPGDLWHLGRHFLLVG 179

Query: 164 DSTNGADVKKLMNGKELALLLLTDPPYNVAYEGKTKDSLTIKNDSMDNDSPRQFLVNAPSS 223
            DST   DVK+LM  E AD++ TDPPYNV YEG T   + IKND+M++  F QFL +AF +
Sbjct: 180 DSTKIEDVKKLMNGNEKALMIPTDPPYNVDYEGAT--GMKIKNDNMEDSEFYQFLFDAFVA 237

Query: 224 ANEVMKPGAVFYIWHADSEGYNFRGACFDIGWTVRQCLIWNKNSMVLGRQDYHWKHEPCL 283
             +V K G   Y+ HADSEG  FR A  D G+ ++QCLIW KNS+VLGRQDYHW+HEP L
Sbjct: 238 MYQVTKEGGPIYVCHADSEGLTFRKAFQDSGFLLKQCLIWVKNSLVLGRQDYHWRHEPIL 297

Query: 284 YGWKDGAGHLWASDRKQTSVID---------------------------------------- 305
           YGWK GA H W   RKQ++VI+
Sbjct: 298 YGWKEGAAHKWYGGRKQSTVIEDPVDLAITPKVDHVLLTFNNGISSTVVKVPSYEIIHDG 357
```

```
Query: 306 ---------YEKPQRNGVHPTMKPVGLFDYQIKNNTKGSDIVLDLPGSSGTTLIACKSNG 356
                    E+P+RN  HPTMKP+ L     I+N++K  + VLD PGSSG+TLIACE  G
Sbjct: 358 SDEGMTTWRIERPKRNADHPTMKPIALCARAIQNSSKPGERVLDPFGSSGSTLIACEQTG 417

Query: 357 RHARLMEYDPKYVDVIIKRWEELTGESVIQL 387
           R   +MEYDP Y +VII+RWEE TG++ ++L
Sbjct: 418 RICHMMEYDPVYAEVIIRRWEEWTGQNAVKL 448
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2572

A DNA sequence (GASx677) was identified in *S.pyogenes* <SEQ ID 7643> which encodes the amino acid sequence <SEQ ID 7644>. Analysis of this protein sequence reveals the following:

```
Possible site: 54

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4744(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2573

A DNA sequence (GASx678) was identified in *S.pyogenes* <SEQ ID 7645> which encodes the amino acid sequence <SEQ ID 7646>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.27   Transmembrane   90 - 106 (  90 - 106)

----- Final Results -----
               bacterial membrane --- Certainty=0.1107(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2574

A DNA sequenee (GASx679) was identified in *S.pyogenes* <SEQ ID 7647> which encodes the amino acid sequence <SEQ ID 7648>. Analysis of this protein sequence reveals the following:

```
Possible site: 19
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.3408(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA66734 GB:X98106 minor capsid protein [Bacteriophage phi3le]
 Identities = 213/494 (43%), Positives = 323/494 (65%), Gaps = 19/494 (3%)

Query: 1   MGVIQKIKNLVTRSKYVM-TTQSLTNITDHPKIAISKLEYDRITTNLKYYKSDWDSVLYL 59
           MG+IQ+IK+L  +       T SL+ ITD P+I+I   EY RI T+L YY      + Y
Sbjct: 1   MGLIQRIKDLFWKGAAATGVTGSLSKITDDPRISIDPDEYVRIQTDLDYYSDKLQYIHYQ 60

Query: 60  NTDGETKKRDLNHLPIARTAAKKIASLVFNEQAEIKV-DDDANNEFISETLKNDRFNNDF 118
            +DG  KKR  N + +A+TAA++IAS+VFNE+AEI V D++ A++P+++ L+++ F   F
Sbjct: 61  ASDGIKKKRLKNTINMAKTAARRIASVVFNEKAEIHVKDNNEADKFLNDVLENDPKNKF 120

Query: 119 ERYLESCLALGGLAMRPYVDGDKVRVAFVQAPVFLPLQSNTQDVSSAAVVIKSVKTINGK 178
            E  LE  +ALGG AMRPY+DG+ +++A+V+A  F PLQSNT D+S AA+  ++ +T + +
Sbjct: 121 EEALEKGVALGGFAMRPYIDGNHIKIAWVRADQFYPLQSNTNDISFAAIASRTQRTESNQ 180

Query: 179 EVYYTLIEFHEWQSSDDYVISNELYRSDDKAKVGSRVPLS--EVYKDLKDEAKVTDVTRP 236
             YYTL+EFH+WQ +  Y I+NELY+SD    VG++VPLS   VYK+L +  ++ + + RP
Sbjct: 181 TKYYTLLEFHQWQDNGSYQITNELYKSDSPDIVGNQVPLSTLPVYKELAPQVTISGLQRP 240

Query: 237 IFTYLKTPQMNNKDINSPLGLSIFDNAKTTIDFINTYDEFMWEVKMGQRRVAVPESLTA 296
            +F Y KTPG NN +I SPLGL + DNAK  +D IN T+D+F+WE+++GQ+ +AV   +
Sbjct: 241 LPAYFKTPGANNINIESPLGLSVVDNAKHVLDDINDTHDQFIWEIRLGQKHIAVQPGMLR 300

Query: 297 LTVRTADGDVVPRPRFESDQNVYIRMGGRDLDSSAIQDLTTPIRADDYIKAINEGLSLFE 356
                D   +P F+++QNVY+ +   D +   ++D+TTPIR   Y  AI+  +  FE
Sbjct: 301 F-------DDEHKPTFDTEQNVYVGVLSDDNNGLGVKDMTTPIRTVQYKDAIDHFIKEFE 353

Query: 357 MQIGVSAGLFSFDGKSMKTATEIVSENSDTYQMRNSIVTLVEQSLKELVISIPEIAKAYD 416
            +QIG+S G FS+    +KTATE+VS NS TYQ R+S +T+VE+++ EL  SIPE+A A
Sbjct: 354 VQIGLSTGTFSYSNDGVKTATEVVSNNSMTYQTRSSYLTMVEKAIDELCQSIFELANAGA 413

Query: 417 LYQSEVP--SMDNISISL------DDGVFTDRDAELDYWIKVVNAGFGTREMAIQKVLNV 468
            L+    P  ++D+ S  L          DDGVF ++D +L+  KV+  G  +++  +Q+  +
Sbjct: 414 LPDDSKPLFTLDSASQPLDIECHPDDGVFVNKDKQLEEDAKVLAIGALSKQTFLQRNYGM 473

Query: 469 TREKAQEIAAEINT 482
            T+E+A E  A+I +
Sbjct: 474 TDEQAAEELAKIQS 487
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2575

A DNA sequence (GASx680) was identified in *S.pyogenes* <SEQ ID 7649> which encodes the amino acid sequence <SEQ ID 7650>. Analysis of this protein sequence reveals the following:

```
Possible site: 48

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.1840(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB53790 GB:AJ242593 gp4 [Bacteriophage A118]
 Identities = 114/385 (29%), Positives = 187/385 (47%), Gaps = 23/385 (5%)

Query: 8   LNDEQLLLEASQLSDMYHQLTLDLFDQVIERIKARGSASLADNFYLWQANKLHDVGLLNA 67
            L   QL L    + D+Y  L  +LF  ++ R+K + + S ADN   WQ  KL+ V  L+
Sbjct: 3   LTPRQLDLLPVQPIVDVYTGLENSLFTLIVRRLKTKKNIS-ADNVLAWQIEKLNQVHALDQ 61

Query: 68  DNIKLIAKYSGIAEAQLRYIIKNEGFKIYKNTSEQLEEALGRESGV-------NSTIQDD 120
              I+ I+K SG++  +L  ++K+ G+   K       +    E+G           TI D
Sbjct: 62  QMIERISKASGVSAKKLFSVVKDAGYSDLKQVDNYFSKLA--EAGAVLPLVSDGQTIVDK 119

Query: 121 LSNYARQAIDDVHNLTNTTLPFSVIGAYQGIIQDAVAGVVTGLKTPDQAINQTVIKWFKK 180
            +   +  +   N T+      Y  II +    V+ GLKT  QA+ +TV K+ +
Sbjct: 120 VMRSYFKLAESNYKRINQTMLSQARQIYSDIIHETTQSVLAGLKTHRQALAETVTKFAEN 179

Query: 181 GPYGFTDKAGRKWRADSYARTVINTTTWRVFNRAKEAPARSPGIDTFYYSKKATAREMCA 240
            G    DKA ++W  ++Y RTV  TT   V+N  ++    E+G+D   S+  AR  C+
Sbjct: 180 GVPALVDKANKRWTPEAYVRTVTRTTVNSVYNSVEDERMNEYGVDLVRISQHVGARPTCS 239

Query: 241 PLQHQIV---TTGEARESGGIKILALSD----YGHGEPDGCLGINCKHTKTPFVVGVNSK 293
             +Q +++   + E R + G K +++      YG+G  DG  G NC+H +  F+ G+N
Sbjct: 240 IVQGKVICLLSVEETRSKVGNKMSIYSPELRYGYG--DGIFGCNCRHHRFAPIBGINIA 297

Query: 294 PELPEHLKNITPAQAKANANAQAKQRAIERSIRKSKELLHVAKQLGDKELIRQYQSDVRS 353
            P+  E   I   + K      +QR +ER IR +K  L  A++LGD+  +++ +  VR+
Sbjct: 298 PDESE---LIDEEENRVYALSQQQRLMERDIRAAKRKLSNAEELGDELAVKKAKQAVRT 354

Query: 354 KQDALNYLINNNAPLHRNQARKRY 378
            KQ  L   +  +  L R  +REK Y
Sbjct: 355 KQSKLRAFVKTHN-LTRQYSREKVY 378
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2576

A DNA sequence (GASx681) was identified in *S.pyogenes* <SEQ ID 7651> which encodes the amino acid sequence <SEQ ID 7652>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2756(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2577

A DNA sequence (GASx682) was identified in *S.pyogenes* <SEQ ID 7653> which encodes the amino acid sequence <SEQ ID 7654>:

```
TLDNQSVIKAIGDTVDYIKKNYKRKWGK
```

Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
            bacterial outside --- Certainty=0.3000(Affirmative) < succ>
           bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
          bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2578

A DNA sequence (GASx683) was identified in *S.pyogenes* <SEQ ID 7655> which encodes the amino acid sequence <SEQ ID 7656>. Analysis of this protein sequence reveals the following:

```
Possible site: 60

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
          bacterial cytoplasm --- Certainty=0.5288(Affirmative) < succ>
           bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2579

A DNA sequence (GASx685) was identified in *S.pyogenes* <SEQ ID 7657> which encodes the amino acid sequence <SEQ ID 7658>:

```
GATEVGANRVVSGVYGEVLGVQIVRSRKCPKGTAYMVRKGALRIMLKRNTMVETDRDTTKAINQIVANKHYGVYLYKAEKAVKITLKDAAK
K
```

Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
          bacterial cytoplasm --- Certainty=0.1750(Affirmative) < succ>
```

```
bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA59185 GB:X84706 major head protein [Bacteriophage B1]
 Identities = 138/270 (51%), Positives = 186/270 (68%), Gaps = 6/270 (2%)

Query: 1   MAVGTTXMAQMLDPEVLADMIDAEVGKAIRFAPLARVDTTLEGQPGTTLTVPK-WDYIGD 59
            M+   T +A +++PEVLA ++  E+ KA+RFAPLA+VDTTL+GQPG TL  P  + YIGD
Sbjct: 1   MSKQKTTLADLVNPEVLATIVSYELMKALRFAPLAQVDTTLQGQPGNTLKFPDPPTYIGD 60

Query: 60  AEDVAEGEAIPWTQLGFKKTIMTIKKAGKGVEITDEAILSGYGDPVGQAAKQIVEAIDHK 119
            A DVAEG  I + ++G    ++TIKKA KG EITDEA LSGYGDP+G++ KQ+  ++ +K
Sbjct: 61  AADVAEGGEISLDKIGTTTKSVTIKKAAKGTEITDEAALSGYGDPIGESNKQLGLSLANK 120

Query: 120 VDADVLDALSKSTQTVSATATVDGVSKALDIFNDEDDAETVIVMNPADASTLRLDAAKEN 179
            VD D+L A   ++QTV   A VDGV  ALDIFNDED    V+++NP DA+ +R DA  +
Sbjct: 121 VDDDLLSAAKTTSQTVSTKANVDGVQAALDIFNDEDAQAYVLIVNPKDAAKIRKDANAKN 180

Query: 180 LGATEVGANRVVSGVYGEVLGVQIVRSRKCPKGTAYMVR----KGALRIMLKRNTMVETD 235
            +G +EVGAN +++G Y +VLG QIVRS+K  +G+A M +       AL+++LKR    VETD
Sbjct: 181 IG-SEVGANALINGTYADVLGAQIVRSKKLAGGSALMFKIVSNSPALKLVLKRGVQVETD 239

Query: 236 RDITKAINQIVANKHYGVVLVKAEKAVKIT 265
            RDI      I A++HY  YLX   K V IT
Sbjct: 240 RDIVTKTTVITADEHYAAYLYDLTKVVNIT 269
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2580

A DNA sequence (GASx686) was identified in *S.pyogenes* <SEQ ID 7659> which encodes the amino acid sequence <SEQ ID 7660>. Analysis of this protein sequence reveals the following:

```
Possible site: 35

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
            bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
           bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2581

A DNA sequence (GASx687) was identified in *S.pyogenes* <SEQ ID 7661> which encodes the amino acid sequence <SEQ ID 7662>. Analysis of this protein sequence reveals the following:

```
Possible site: 54

>>> Seems to have no N-terminal signal sequence
```

```
----- Final Results -----
                bacterial cytoplasm --- Certainty=0.2942(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2582

A DNA sequence (GASx688) was identified in *S.pyogenes* <SEQ ID 7663> which encodes the amino acid sequence <SEQ ID 7664>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.2844(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC00538 GB:L02496 unknown protein [Bacteriophage LL-H]
 Identities = 35/86 (40%), Positives = 48/86 (55%), Gaps = 6/86 (6%)

Query: 24  KLIMNNQVMMSMNPYVPYRDGALRGSSRANSVGVTWSGPHARAQFYGGAYNKYKSFKFKK 83
           +L + NQ+    M  YVP R G LR  S  N  G+ ++  +ARAQFYG           + +
Sbjct: 20  RLQVLNQMHQDMEQYVPKRAGFLRSQSFVNDTGIHYTAYARAQFYGFV----NGHRVRN 75

Query: 84  YTTPGTGKRWDKRALANATIVKDWEK 109
           Y+TPGTG+RWD +  A A    DW+K
Sbjct: 76  YSTPGTGRRWDLK--AKAVYKADWQK 99
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2583

A DNA sequence (GASx689) was identified in *S.pyogenes* <SEQ ID 7665> which encodes the amino acid sequence <SEQ ID 7666>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.2892(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA66741 GB:X98106 minor capsid protein [Bacteriophage phig1e]
```

```
 Identities = 36/109 (33%), Positives = 64/109 (58%), Gaps = 2/109 (1%)

Query: 17  DLGIKPRLDYLTRQEDLAIYPMPGGKVNNEYMDGTREISLPFEIAIKTKNQELASTVMWT 76
           +L +K  L YLT  + L++YP+PG +V +E   G ++  + +E+ ++TKNQ+ A+T +W
Sbjct: 16  NLPMKCTLGYLTAADSLSLYPLPGSRVLDKDYAGNQQWQMNYEVGMKTKNQQQANTTLWL 75

Query: 77  INSALSNPDL-KLPSLNHSYTFISLDVE-KPFLNDLSDQGFYIYVLDIT 123
           ++ AL       L S N S+ F SL +  +P +++    QG+  Y L  +
Sbjct: 76  VSQALDVLTADDLVSSNGSFEPESLTINGQPSISEQDTQGYSTYQLSFS 124
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2584

A DNA sequence (GASx690) was identified in *S.pyogenes* <SEQ ID 7667> which encodes the amino acid sequence <SEQ ID 7668>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1626(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB53798 GB:AJ242593 major tail shaft protein [Bacteriophage
          A118]
 Identities = 54/133 (40%), Positives = 77/133 (57%), Gaps = 9/133 (6%)

Query: 1   MRQKNALRGHPIAPYVKGEEKTEVTKEKLLELARWIKDISDDTDEKTSDEAYYDGDGTEE 60
           MR KNA   + +A  V G  + +  +      L++WI ++SDD  + TE++  YDGDG E+
Sbjct: 1   MRIKNAKTKYSVAEIVAGAGEPDWKR------LSKWITNVSDDGSDNTEEQGDYDGDGNEK 55

Query: 61  TTVVGVKGAYTFEGTYDPEDKAQAHIASLKYKLGDERKVWHLIVSADGKTQWLGVATVTE 120
           T V+G   AYTFEGT+D ED+AQ  I + K +  + R +   I   D +T  +G ATV+E
Sbjct: 56  TVVLGYSEAYTFEGTHDREDEAQNLIVA-KRRTPENRSIMFKIEIPDTETA-IGKATVSE 113

Query: 121 I--IAGSGAAARF 131
           I   AG G A  F
Sbjct: 114 IKGSAGGGDAIEF 126
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2585

A DNA sequence (GASx691) was identified in *S.pyogenes* <SEQ ID 7669> which encodes the amino acid sequence <SEQ ID 7670>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3521(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2586

A DNA sequence (GASx692) was identified in *S.pyogenes* <SEQ ID 7671> which encodes the amino acid sequence <SEQ ID 7672>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3438(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB53801 GB:AJ242593 gp15 [Bacteriophage A118]
 Identities = 67/191 (35%), Positives = 110/191 (57%), Gaps = 17/191 (8%)

Query: 11  FEFRGSIYPIDLSFNKVLDVFDVIDDDFLNEAEKCFLCLDILLDRTDLPFTYAVD----- 65
           +E+ G+ Y +DL+F+ VL V D+ +D+ L++  +  L +D+L    D+P+  + +
Sbjct: 12  YEYSGKEYKLDLAFDNVLKRVIDLTEDNSLSDVFRANLAIDVLF-ADDMPWPRSNEEDEYA 70

Query: 66  -------LNVYIKTNFIDAERPEKPQLDIKGNPMPVVKEKEDNKKVI---DLSLDAEFIY 115
                + + I TN+I  E  +    DI GN MP      D+ + I    L+ DA++IY
Sbjct: 71  NIEEKSLVLIDIFTNYIVKENDDGLLYDIDGNKMPSATNNNDDAERIASYSLTQDADYIY 130

Query: 116 ASFRQAYQINLLKEQNRLSWIEFKALLNALPDDTVMQRIIAIRQWE-DDGEGSKKYRDNM 174
           ASF Q Y I+LL  + ++ W +F+ALL +L DDT ++ II IRQ E   G+G++K R+ +
Sbjct: 131 ASFLQDYNIDLLDSRGKMHWYKPRALLESLRDDTTIKTIIGIRQAELPSGKGTEKERNSL 190

Query: 175 RKLKAKYSLDE 185
            KLK +Y L +
Sbjct: 191 IKLKNRYKLKD 201
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2587

A DNA sequence (GASx694) was identified in *S.pyogenes* <SEQ ID 7673> which encodes the amino acid sequence <SEQ ID 7674>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4143(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAG18639 GB:AY007505 unknown [Streptococcus mitis]
 Identities = 48/157 (30%), Positives = 85/157 (53%), Gaps = 10/157 (6%)

Query: 86  DLELSWEPDYIYKATHTTPPSIKEVLRNPGRLKINFLIHPIKYLKTGKQEVPLVNG-GTL 144
           +LE S+ P+ ++ A H    S K    +  +LKI    + P +Y KT   E    NG GT+
Sbjct: 81  ELEPSYHPESVFYA-HFLTASYKPFGNHAWQLKIKLNMQPPRYQKTVNPES--YNGPGTI 137

Query: 145 QNPGNVQAKPILKIKGTGNGILTINDPSTGLENVQSELVIDMERHLVYKDVLSAWDNIVR 204
            NPG + ++PI+++G G+  +TI   ET   NV+++  ID +    +++ +A   +
Sbjct: 138 NNPGTIYSEPIEVQGDGDVSITIGR-ETMYLNVKTKATIDCRQG--RQNIYNATGAVQN 194

Query: 205 TERHRMPLFDV--GQNKISWTGS-FTITTAVPNWGVKV 238
            T R R   F++  G++ I++TG+   +   PNW  K+
Sbjct: 195 TLRKRGGFFEIPTGRSGITFTGNVLRLIIRPNWRYKI 231
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2588

A DNA sequence (GASx695R) was identified in *S.pyogenes* <SEQ ID 7675> which encodes the amino acid sequence <SEQ ID 7676>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -2.60   Transmembrane   15 -  31 (  15 -  31)

----- Final Results -----
                bacterial membrane --- Certainty=0.2041(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2589

A DNA sequence (GASx697) was identified in *S.pyogenes* <SEQ ID 7677> which encodes the amino acid sequence <SEQ ID 7678>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.3348(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA86895 GB:U28144 hyaluronidase [Streptococcus pyogenes]
 Identities = 326/337 (96%), Positives = 329/337 (96%)
```

```
Query: 1    MSENIPLRVQFKRMKAAEWARSDVILLESEIGFEITVTGFARAGDGHNRFSDLGYISPLDY 60
            MSENIPLRVQFKRMKAAEWARSDVILLESEIGFEITVTGFARAGDGHNRFSDLGYISPLDY
Sbjct: 1    MSENIPLRVQFKRMKAAEWARSDVILLESEIGFEITVTGFARAGDGHNRFSDLGYISPLDY 60

Query: 61   NLLTNKPNIDGLATKVETAQKLQQKADKETVYTKAESKQELDKKLNLKGGVMTGQLKFKP 120
            NLLTNKPNIDGLATKVETAQKLQQKADKETVYTKAESKQELDKKLNLKGGVMTGQLKFKP
Sbjct: 61   NLLTNKPNIDGLATKVETAQKLQQKADKETVYTKAESKQELDKKLNLKGGVMTGQLKFKP 120

Query: 121  AATVAYSSSTGGAVNIDLSSTRGAGVVVYSDNDTSDGPLMSLRTGKETFNQSALFVDYKG 180
            AATVAYSSSTGGAVNIDLSSTRGAGVVVYSDNDTSDGPLMSLRTGKETFNQSALFVDYKG
Sbjct: 121  AATVAYSSSTGGAVNIDLSSTRGAGVVVYSDNDTSDGPLMSLRTGKETFNQSALFVDYKG 180

Query: 181  TTNAVNIAMRQPTTPNFSSALNITSGNENGSAMQLRGSEKALGTLKITHENPSIGADYDK 240
            TTNAVNIAMR  TTPNFSSALNITSGNENGSAMQLRGSEKALGTLKITHENPSIGADYDK
Sbjct: 181  TTNAVNIAMRHATTPNFSSALNITSGNENGSAMQLRGSEKALGTLKITHENPSIGADYDK 240

Query: 241  NAAALSIDIVKKTNGAGTAAQGIYINSTSGTTGKLLRIRNLSDDKPYVKSDGGFYAKETS 300
            NAA    + + K+ NGAGTAAQGIYINSTSGTTGKLLRIRNLSDDKPYVKSDGGFYAKETS
Sbjct: 241  NAARYPLILSKRQNGAGTAAQGIYINSTSGTTGKLLRIRNLSDDKPYVKSDGGFYAKETS 300

Query: 301  QIDGNLKLKDPTANDHAATKAYVDKAISELKKLILKK 337
            QIDGNLKLKDPTANDHAATKAYVDKAISELKKLILKK
Sbjct: 301  QIDGNLKLKDPTANDHAATKAYVDKAISELKKLILKK 337
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2590

A DNA sequence (GASx698) was identified in *S.pyogenes* <SEQ ID 7679> which encodes the amino acid sequence <SEQ ID 7680>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4208(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>

RGD motif 54-56
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA98102 GB:M19348 ORF [Streptococcus pyogenes phage H4489A]
 Identities = 250/648 (38%), Positives = 351/648 (53%), Gaps = 75/648 (11%)

Query: 1    MSRDPTLILDESNLVIGKDGRVHYTFPTEDDNPKVRLASKCLGTAHFNQLMIERGDQATS 60
             MSRDPT  ++E +L    DGR + TF  +  + VRL S CLG     +L +E  +
Sbjct: 1    MSRDPTYTINEHDLSFA-DGRFYVTFRADKSSETVRLNSSCLGNTIIKKLQVEDDNTMHD 59

Query: 61   YVAPVVVEGTGNPTGLPKDLKKISLELTDTANSQLWSKIKLTNRGMLQEYYDGKIKTRIV 120
             +V P V   T   GL + +KK+ L+L D   S LW KIK  N+ ML EY + ++ + I
Sbjct: 60   FVKPKVT--TQQAPGLAQCVKSLDLQLKDP-KSDLWGKIKPNNKAMLVEYANKSMSSAIA 116

Query: 121  NSARGVATRISEDTDKKLALINDTIDGIRREYRDDADRKLSASYQAGIEGLKATMNNDKIG 180
              SA  +  ++    D++ +    T++GI++  +
Sbjct: 117  QSAEQILLQVKSIDDERYSKFEQTLNGIKQTVKSES---------- 152

Query: 181  LQAETKASAQGLSQKYDDELRKLSAKITTTSSGTTEAYESKLAQLRAEFTRSNQGTRTEL 240
             ++++   L+  +D  +  L  K +  S  T ++  S+L           G  + L
Sbjct: 153  ----VESARTQLASMPDSRISGLDGKYSRLSQ-TIDSLSSRLD--------DGVGNYSTL 199

Query: 241  ESQISGLRAVQQSTASQISQEIRDRDGAVSRVQQSLESYQRRMQDAEENYSSLTHTVRGL 300
```

```
              ++SG              I   + +    VSR+ Q+ +  Q ++ +A +NYSSL+ TV+GL
Sbjct: 200 SQKVSG----------IDLRVSNAANDVSRLSQTAQGLQSQITNANQNYSSLSQTVQGL 248

Query: 301 QSDVGSPTGKIQSRLTQLAQQIEQRVTRDGVMSIISGAGDGIKLAIQKAGGINAKMSGNE 360
            Q+ V          SR+ QL+  I  +VT+  V + I+ + D I  AI+      + KK+G+E
Sbjct: 249 QTTVRDNQSNATSRINQLSDLISTKVTKGDVETTIAQSYDKIAFAIRDKLPAS-KMTGSE 307

Query: 361 IISAINLNSYGVTIAGKHIALDGNTTVNGTFTTKIARATIKRADQIIAGTIDAARIRVIN 420
            IISAINL+  GV I GK+I LDGN+ ++      K A     + A +I  G ++A+RI
Sbjct: 308 IISAINLDRSGVKITGKNITLDGNSYISNA-VIKDAHIANMDAGKINTGYLNASRIAAEA 366

Query: 421 LNASSIVGLDANFIK--AKIGY----------------AIT---DLLEGKVIKARNGAMLI 460
            +    I    A F K  A  GY                A+T     + G V+ A NGA
Sbjct: 367 ITGDKIKMDYAFFNKLITANSGYFRTLFAKNIFTTSVQAVITSASKITGGVLGATNGASRW 426

Query: 461 DLNTAKNDFNSDATINFNSKNNALVRKDGTWTAFVHFSNATPKGYTGSAIYASIGITSSG 520
            DLN+A +DFN DATINFNSKNNALVRK GT+TAFVHFSNATPKGY GSALYASIGITSSG
Sbjct: 427 DLNSANIDFNRDATINFNSKNNALVRKSGTNTAFVHFSNATPKGYRGSALYASIGITSSG 486

Query: 521 DGVNSASSGRFAGLRSFRYATGYNHTAAVDQTEIYGDNVLVVDDPNITRGFKFRPDKMQK 580
            DG++SASSGRF G+R FRYA G  HTA VDQ EIYGD+++   DDFNI RGFK RP  M K
Sbjct: 487 DGIDSASSGRFCGVRPFRYAEGLQHTAKVDQAEIYGDDIVFSDDFNIDRGFKMRPSLMPK 546

Query: 581 MLDMNDLYAAVVALGRCWGHLANVGWNTAHSNFTSAVNRELNNYITKI 628
            M+D+N +Y A++ALGRCW H  N  W+  + +  SA+  E N +I  +
Sbjct: 547 MVDLNKMYQALLALGRCWLHANNTAWSW-NFDTRSAIIAEYNAHINMI 593
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2591

A DNA sequence (GASx699) was identified in *S.pyogenes* <SEQ ID 7681> which encodes the amino acid sequence <SEQ ID 7682>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3323(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2592

A DNA sequence (GASx701) was identified in *S.pyogenes* <SEQ ID 7683> which encodes the amino acid sequence <SEQ ID 7684>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1017(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2593

A DNA sequence (GASx702) was identified in *S.pyogenes* <SEQ ID 7685> which encodes the amino acid sequence <SEQ ID 7686>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -3.03   Transmembrane    2 -  18 (   1 -  23)

----- Final Results -----
                bacterial membrane --- Certainty=0.2211(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2594

A DNA sequence (GASx703) was identified in *S.pyogenes* <SEQ ID 7687> which encodes the amino acid sequence <SEQ ID 7688>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -3.45   Transmembrane   36 -  52 (  36 -  55)

----- Final Results -----
                bacterial membrane --- Certainty=0.2381(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC39287 GB:AF115103 orf87 gp [Streptococcus thermophilus
          bacteriophage Sfi21]
 Identities = 43/73 (58%), Positives = 61/73 (82%)

Query: 1  MINLKLRLQNKVTLMAILGAIFLLAQQLGIKLPSNIADIANIAVTLLVLLGVVTDPTTKG 60
          MIN KLRLQNK TL+A++ A+FL+ QQ G+ +P+NI +  NT V +LV+LG++TDPTTKG
Sbjct: 8  MINFKLRLQNKATLVALISAVFLMLQQPGLHVPNNIQEGINTLVGILVILGIITDPTTKG 67

Query: 61 LSDSEQALTYHEP 73
          ++DSE+AL+Y +P
Sbjct: 68 IADSERALSYIQP 80
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2595

A DNA sequence (GASx707R) was identified in *S.pyogenes* <SEQ ID 7689> which encodes the amino acid sequence <SEQ ID 7690>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-10.35   Transmembrane    9 -  25 (   1 -  27)

----- Final Results -----
                bacterial membrane --- Certainty=0.5140(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2596

A DNA sequence (GASx714R) was identified in *S.pyogenes* <SEQ ID 7691> which encodes the amino acid sequence <SEQ ID 7692>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.1401(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2597

A DNA sequence (GASx715) was identified in *S.pyogenes* <SEQ ID 7693> which encodes the amino acid sequence <SEQ ID 7694>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.0417(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2598

A DNA sequence (GASx726) was identified in *S.pyogenes* <SEQ ID 7695> which encodes the amino acid sequence <SEQ ID 7696>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.17   Transmembrane   18 -  34 (  18 -  35)

----- Final Results -----
                bacterial membrane --- Certainty=0.1468(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2599

A DNA sequence (GASx728R) was identified in *S.pyogenes* <SEQ ID 7697> which encodes the amino acid sequence <SEQ ID 7698>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1795(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF61314 GB:U96166 unknown [Streptococcus cristatus]
 Identities = 149/194 (76%), Positives = 162/194 (82%)

Query: 1   LSAIIRQSTSKRISDKRGIYLVEKLVSLAKQSYFTVTKTSPMIEEVRYYAKELLRLSERR 60
           L  IIRQSTSKRIS+KR  YL +KL+ LAKQS+  V KTSPM+EEVRYYA+ELLRLSERR
Sbjct: 56  LYEIIRQSTSKRISEKRIAYLTDKLIKLAKQSFCAVKKTSPMLEEVRYYAQELLRLSERR 115

Query: 61  QAIFDKMVASAQPLPEDKILRSIPSIVETTATSIIGELGAIRRFQSANQINAFIGIDFRH 120
           Q + + MVA AQPLPE  ILRSIP I ETTATSIIGELG I RFQS NQ NAFIGID RH
Sbjct: 116 QVVLNDMVAIAQPLPEYDILRSIPGIAETTATSIIGELGDIHRFQSTNQFNAFIGIDLRH 175

Query: 121 YESGNYLAQEHITKRGNPYAPKILFKCIHDIAFASHTNPCHIADFYEKRKRQSQTASTKP 180
           YES N+LA+EHITKRGNPYA KILFKCIH+IA ASHTNPCHIADFYEKRKRQS  ASTKP
Sbjct: 176 YESRNFLAKEHITKRGNPYARKILFKCIHNIASASHTNPCHIADFYEKRKRQSTIASTKP 235

Query: 181 HTIASRHCLVRQCF 194
           TIAS H L+R  +
Sbjct: 236 LTIASIHRLIRTMY 249
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2600

A DNA sequence (GASx729R) was identified in *S.pyogenes* <SEQ ID 7699> which encodes the amino acid sequence <SEQ ID 7700>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2363(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2601

A DNA sequence (GASx730R) was identified in *S.pyogenes* <SEQ ID 7701> which encodes the amino acid sequence <SEQ ID 7702>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2602

A DNA sequence (GASx734) was identified in *S.pyogenes* <SEQ ID 7703> which encodes the amino acid sequence <SEQ ID 7704>. Analysis of this protein sequence reveals the following:

```
Possible site: 52

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4001(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2603

A DNA sequence (GASx735) was identified in *S.pyogenes* <SEQ ID 7705> which encodes the amino acid sequence <SEQ ID 7706>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -3.66   Transmembrane  276 - 292 ( 274 - 292)

----- Final Results -----
                bacterial membrane --- Certainty=0.2466(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2604

A DNA sequence (GASx736) was identified in *S.pyogenes* <SEQ ID 7707> which encodes the amino acid sequence <SEQ ID 7708>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3998(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2605

A DNA sequence (GASx737) was identified in *S.pyogenes* <SEQ ID 7709> which encodes the amino acid sequence <SEQ ID 7710>. Analysis of this protein sequence reveals the following:

```
Possible site: 60

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-12.74   Transmembrane   77 -  93 (  69 -  99)
   INTEGRAL    Likelihood = -4.14   Transmembrane  152 - 168 ( 151 - 170)
   INTEGRAL    Likelihood = -1.17   Transmembrane  196 - 212 ( 194 - 212)

----- Final Results -----
                bacterial membrane --- Certainty=0.6095(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2606

A DNA sequence (GASx738) was identified in *S.pyogenes* <SEQ ID 7711> which encodes the amino acid sequence <SEQ ID 7712>. Analysis of this protein sequence reveals the following:

```
Possible site: 37

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-13.16   Transmembrane    44 -  60 (  39 -  71)
   INTEGRAL    Likelihood =-10.24   Transmembrane    94 - 110 (  81 - 114)
   INTEGRAL    Likelihood = -7.64   Transmembrane   185 - 201 ( 179 - 207)
   INTEGRAL    Likelihood = -7.48   Transmembrane   132 - 148 ( 130 - 158)
   INTEGRAL    Likelihood = -2.76   Transmembrane   208 - 224 ( 204 - 225)
   INTEGRAL    Likelihood = -0.06   Transmembrane   153 - 169 ( 152 - 169)

----- Final Results -----
                bacterial membrane --- Certainty=0.6265(Affirmative) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2607

A DNA sequence (GASx742) was identified in *S.pyogenes* <SEQ ID 7713> which encodes the amino acid sequence <SEQ ID 7714>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -7.80   Transmembrane   887 - 903 ( 882 - 906)
   INTEGRAL    Likelihood = -4.88   Transmembrane     6 -  22 (   5 -  23)

----- Final Results -----
                bacterial membrane --- Certainty=0.4121(Affirmative) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>

LPXTG motif: 877-881
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB46409 GB:AL096743 putative large secreted protein
          [Streptomyces coelicolor A3(2)]
 Identities = 231/599 (38%), Positives = 329/599 (54%), Gaps = 43/599 (7%)

Query: 278 TSSNSDASSRNIVKIGRIQGASHTSPLLKKAVTVEQVVVTYL---DDSTHFYVQDLNGDG 334
           T +++ +++   V+I ++QG++  SP   + VT     +VT +     S  F++QD   D
Sbjct: 28  TPAHAASAAAGPVRIHDVQGSTRLSPYAGEQVTDVAGIVTGVRGYGSSKGFWMQDPLPDA 87
```

```
Query: 335 DLATSDGIRVPAKNA-KVQVGDVLTISGEVEEPFGRGYEERKQTDLTITQIVAKAVTK-T 392
           D ATS+G+ VP    A +V VGD +T+SG V E+    G      Q+    +T+I     VT   +
Sbjct: 88  DPATSEGVFVPTSRAPEVAVGDAVTVSGTVSEYVPGGTSSGNQS---LTEITRPTVTVVS 144

Query: 393 GTAQVPSPLVLGKDRIAPANIIDNDGLR-------VFDPEEDAIDVWESMEGMLVAVDDA 445
            G   +P+   +    + A   + DG             P   A+DV+ES+EGM V V DA
Sbjct: 145 GGNAIPAATTVSARSVPRAYAPEGDGAANGSVNALPLRPGTVALDYYESLEGMNVRVADA 204

Query: 446 KILGPMKN-KEIYVLPGSSTRPLNNSGGVLLPANSYNIDVIPVLFKKGKQI----IKAGD 500
            +++G       E++V       P    G V     + NT  + +     GK          GD
Sbjct: 205 RVVGASDPYTSLWVTVKPWENPNRRGGTVYGSYDDQNTGRLQIQ-SLGKPADFPAADVGD 263

Query: 501 SYKGRLAGPVSYS-YGNYKVFVDDSKNMPSLMDGHLKPEKTNLQKDLSKLSIASYNIENF 559
            +  G  AGP+ Y+ YG Y +   +   + +L  G  + E T  Q     +L++A+YN+EN
Sbjct: 264 TLAGTTAGPLDYNQYGGYTLVASE---IGALSSGGTSRSSTRRQS-ARSLAVATYNVENL 319

Query: 560 SANPSSTKDEKVKRIAESFIHDLNAPDIIGLIEVQDNNGPTDDGTTDATQSAQRLIDAIK 619
             +PS   D+     AE+ +H L +PDI+ L E+QDNNG TDDGT   A  +  RLIDAI
Sbjct: 320 --DPS---DDTFTAHAETIVHRLKSPDIVSLSEIQDNNGATDDGTVAADATVGRLIDAIV 374

Query: 620 KLGGPTYRYVDIAPENVDGGQPGGNIRTGFLYQPERVSLSDKPKGGARDA--LTWVNGE 677
             GGP Y +  I P +  DGGQPGGNIR   FL+  PERVS +D+  G A  A  +  V G+
Sbjct: 375 AAGGPRYDWRGIDPVDKADGGQPGGNIRQAFLFNPERVSFTDRAGGDATTATGVRKVRGK 434

Query: 678 --LNLSVGRIDPTNAAWKDVRKSLAAEFIFQGRKVVVVANHLNSKRGDNALYGCVQPVTF 735
             L  S GR+DP N AW+D RK LA EF+F+GR  V VVANH NSK GD  L     QP +
Sbjct: 435 AALTRSPGRVDPANEAWEDSRKPLVGEFVFRGRTVFVVANHFNSKGGDQQLTAQYQPPSR 494

Query: 736 KSEQRRHVLANMLAQFAKE--GAKHQANIVMLGDFNDFEFTKTIQLIE-EGDMVNLVSRH 792
            SE +RH  A ++  F KE   A+  A++V LGD NDFEF++T +++E +G + + V
Sbjct: 495 GSETQRHAQAKVVNTFVKEILAAQKNADVVALGDINDFEFSRTARLLEGDGALWSAVKSL 554

Query: 793 DISDRYSYPHQGNNQTLDNILVSRHLL--DHYEFDMVHVNSPFMEAHGRASDHDPLLLQ 849
             S+RYSY +QGN+Q LD ILVS  +    H  +D VHVN+ F   H + SDHDP +L+
Sbjct: 555 PRSERYSFVYQGNSQVLDQILVSPSVRRGGHLSYDSVHVNAEF---HDQISDHDPQVLR 610
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2608

A DNA sequence (GASx743) was identified in *S.pyogenes* <SEQ ID 7715> which encodes the amino acid sequence <SEQ ID 7716>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2437(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2609

A DNA sequence (GASx756) was identified in *S.pyogenes* <SEQ ID 7717> which encodes the amino acid sequence <SEQ ID 7718>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -4.30   Transmembrane   10 -  26 (   8 -  27)
   INTEGRAL    Likelihood = -3.08   Transmembrane   51 -  67 (  50 -  67)

----- Final Results -----
               bacterial membrane --- Certainty=0.2720(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2610

A repeated DNA sequence (GASx758) was identified in *S.pyogenes* <SEQ ID 7719> which encodes the amino acid sequence <SEQ ID 7720>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA38133 GB:X54225 7 kDa protein [Streptococcus pneumoniae]
 Identities = 31/61 (50%), Positives = 41/61 (66%)

Query: 1  MTNGLKYVLEQMLLFIIAALACLFLAIGLMIGYSFMGDQQSPWHILSMDKWAELVNKFT 60
          M     YV+++++LL+ I+  L  L L IGLM+GY  +G GQ PW ILS  KW EL++KFT
Sbjct: 3  MNKKSSYVVKRLLLVIIVLILGTLALGIGLMVGYGILGKGQDPWAILSPAKWQELIHKFT 62

Query: 61 G 61
          G
Sbjct: 63 G 63
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2611

A DNA sequence (GASx764) was identified in *S.pyogenes* <SEQ ID 7721> which encodes the amino acid sequence.<SEQ ID 7722>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence
```

```
   INTEGRAL    Likelihood = -3.98   Transmembrane   47 -  63 (  46 -  67)

----- Final Results -----
                bacterial membrane --- Certainty=0.2593(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

A related sequence was also identified in GAS <SEQ ID 9149> which encodes the amino acid sequence <SEQ ID 9150>. Analysis of this protein sequence reveals the following:

```
       Possible site: 53
>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -3.98   Transmembrane   35 -  51 (  34 -  55)

----- Final Results -----
                bacterial membrane --- Certainty=0.2593(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2612

A DNA sequence (GASx783) was identified in *S.pyogenes* <SEQ ID 7723> which encodes the amino acid sequence <SEQ ID 7724>. Analysis of this protein sequence reveals the following:

```
Possible site: 43

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-13.16   Transmembrane  142 - 158 ( 132 - 167)
   INTEGRAL    Likelihood =-12.26   Transmembrane  113 - 129 ( 101 - 140)
   INTEGRAL    Likelihood =-10.24   Transmembrane  238 - 254 ( 233 - 260)
   INTEGRAL    Likelihood = -2.76   Transmembrane   34 -  50 (  34 -  51)

----- Final Results -----
                bacterial membrane --- Certainty=0.6265(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA32091 GB:AB010970 ABC-transporter [Streptococcus mutans]
 Identities = 173/269 (64%), Positives = 214/269 (79%), Gaps = 2/269 (0%)

Query: 1   MNFLTKKNRILLREMVKTDFKLRYQGSAIGYLWSILKPLMMFTIMYLVFIRFLRLGGNVP 60
           M+F ++KNRILL+E++KTDFKLRYQGSAIGYLWSILKPLM+F IMY+VF+RFL LGG+VP
Sbjct: 1   MDFFSRKNRILLKELIKTDFKLRYQGSAIGYLWSILKPLMLFAIMYIVFVRFLPLGGDVP 60

Query: 61  HFPVALLLANVIWSFFSEATSMGMVSIVSRGDLLRKLNFSKHIIVFSAVLGALINFLINL 120
           H+PVALLL NVIW+FF E T MGMVS+V+RGDLLRKLNFSK  IVFSAV GA INF IN+
Sbjct: 61  HWPVALLLGNVIWTFFQETTMMGMVSVVTRGDLLRKLNFSKQTIVFSAVSGAAINFGINV 120

Query: 121 VVVLIFALINGVTIS--GYAYLSLFLFIELVLVLVLGIALLLSNVFVYYRDLAQVWEVLLQ 178
           +VVLIFAL+NGVT +     +L + LF+EL++   GIA +LS ++V YRD+  VWEV+LQ
Sbjct: 121 IVVLIFALLNGVTFTPFWNLFLLIPLFLELLLSFSTGIAFILSTLYVRYRDIGPVWEVILQ 180

Query: 179 AGMYATPIIYPITFVLDGHPLAAKLLMLNPVAQMIQDFRYLLIDRANVTIWQMSTNWFYI 238
```

```
              G Y TPIIY +T++      + AKLL+L+P+AQ+IQD R+++LID ANVTIWQM   +
Sbjct: 181 GGFYGTPIIYSLTYIATRSVVGAKLLLLSPIAQIIQDMRHILIDPANVTIWQMINHKSIA 240

Query: 239 VIPVLVPFVILFIGIFVFKKNADRFAEII 267
           VIPVLVP  +  IG  VF  NA +FAEII
Sbjct: 241 VIPVLVPIFVFIIGFLVFNYNAKKFAKII 269
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2613

A DNA sequence (GASx786) was identified in *S.pyogenes* <SEQ ID 7725> which encodes the amino acid sequence <SEQ ID 7726>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3828(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA32094 GB:AB010970 rgpFc [Streptococcus mutans]
 Identities = 381/582 (65%), Positives = 475/582 (81%), Gaps = 1/582 (0%)

Query: 1   MNRILLYVHFNKYNKISAHVTYQLEQMRSLFSKIVFISNSKVSHEDLKRLKQNHCLIDEFL 60
           M R+LLYVHFNKYN+++SHV YQL QMRSLFSK++FISNS+V+  D+K L+   LID+F+
Sbjct: 1   MKRLLLYVHFNKYNRVSSHVVYQLTQMRSLFSKVIFISNSQVADADVKMLREKELIDDFI 60

Query: 61  QRKNKGFDFSAWHDGLIIMGFDKLEEFDSLTIMMDTCPGPIWEMAPYFENFEEKETVDFW 120
           QR+N GFDF+AW DG++ +GFD+L  +DS+T MNDTCPGF+WEM    ++ FE K TVDFW
Sbjct: 61  QRQNSGFDFAAWRDGMVFVGFDELVTYDSVTTMMDTCPGPLWEMYSIYQEPETKTTVDFW 120

Query: 121 GITNNRGTKAPKEHVQSYFMTFKNQVIQNKVPQQFWQSIIEYENVQEVIQHYETQLTSIL 180
           G+TNNR TK+P+EH+QSYF++FK  V+++  F+ FW++I EY++VQ+VI  YET++T+ L
Sbjct: 121 GLTNNRATKSPREHIQSYFISFKASVLRSTAFRDFWENIKEYQDVQKVIDQYETKVTTTL 180

Query: 181 LNEGFSYQTVPDTRKAESSFWPHPDFSYYNPTALLGHVDFIKVKAIDANQMIAFVLLNL 240
           L+ GF Y  VFDT K ++S M H DFSYYNPTAIL H VPFIKVKAID NQMI FVLLN
Sbjct: 181 LDAGFQYDVVFDTTKEDASHMLHADFSYYNPTALLNHRVPFIKVKAIDNNQHITFVLLND 240

Query: 241 IRETTNYPIDLIVSHMSQISLPDTKYLLSQKYLNCQRLAKQTCQKVAVHLHVFYVDLLDE 300
           I++ + YPIDLIVSHMS+I+ PD  YLL  KY+  +       QKVAVHLHVFYVDLL+E
Sbjct: 241 IQKNSTYPIDLIVSHMSEINYPDFSYLLGHKYVKKRSRVDLKNQKVAVHLHVFYVDLLKE 300

Query: 301 FLTAFENWNPHYDLFITTDSDIKRKEIKRILQRKGKTADIRVTGNRGRDIYPMLLLKDKL 360
           FLTAF+ ++F YDLFITTDSD K+ EI+RIL   G+ A + VTGN GRD+ PML LK+ L
Sbjct: 301 FLTAFKQFHFSYDLFITTDSDDKKAEIERILSANCQEAQVFVTGNIGRDVLPMLKLKNYL 360

Query: 361 SQYDYIGHFHYKKSKEADFWAGESSWRKSLIDMLVKPADSILSAFETD-DIGIIIADIPSF 419
           S YD++GHFHYKKSKEADFWAG+SWR+SLIDMLVKPADI+LA  + +  IG++IAD+P+F
Sbjct: 361 SAYDFVGHFHYKKSKEADFWAGQSWREELIDMLVKPADNILAQLQQNPKIGLVIADMPTF 420

Query: 420 RFNKIVNAWNEHLIAQEMMSLWRKMDVKKQIDPQAMDFFVMSYGTFVWFKYDALKSLFD 479
           FR+NKIV+AWNEHLIA EM +LW+KM + K+IDF A  TPVMSYGTFVWFKYDALK LFD
Sbjct: 421 FRYNKIVDAWNEHLIAPEMNTLWQKMMTKKIDFNAPHTPVMSYGTFVWFKYDALKPLFD 480

Query: 480 LELTQNDIPSPPLPQNSILHAIRRLLVYIAWGDSYDFRIVKNPYELTPFIDNKLLNLRED 539
           L LT +D+P EPLPQNSILHAIRRLLVYIAW + YDFRI KNP +LTPFIDNKLLN R +
Sbjct: 481 LNLTDDDVPEEPLPQNSILHAIRRLLIYIAWNEHYDFRISKNPVDLTPFIDNKLLNERGN 540
```

```
Query: 540 EGAHTYVNFNQMGGIKGALKYIIVGPAKAMKYIFLRLMEKLK 581
              +T+V+FN MGGIKGA KYI +GPA+A+KYI  R ++K+K
Sbjct: 541 SAPNTFVDFNYMGGIKGAFKYIPIGPARAVKYILKRSLQKIK 582
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2614

A DNA sequence (GASx787) was identified in *S.pyogenes* <SEQ ID 7727> which encodes the amino acid sequence <SEQ ID 7728>. Analysis of this protein sequence revealed the following:

```
Possible site: 33

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-15.66   Transmembrane  202 - 218 ( 191 - 224)
   INTEGRAL    Likelihood =-10.03   Transmembrane  340 - 356 ( 335 - 365)
   INTEGRAL    Likelihood = -9.08   Transmembrane  270 - 286 ( 263 - 289)
   INTEGRAL    Likelihood = -8.60   Transmembrane  124 - 140 ( 118 - 145)
   INTEGRAL    Likelihood = -4.94   Transmembrane  377 - 393 ( 375 - 395)
   INTEGRAL    Likelihood = -3.29   Transmembrane  291 - 307 ( 290 - 311)
   INTEGRAL    Likelihood = -2.87   Transmembrane  160 - 176 ( 159 - 180)
   INTEGRAL    Likelihood = -2.66   Transmembrane   50 -  66 (  48 -  66)
   INTEGRAL    Likelihood = -1.28   Transmembrane   77 -  93 (  76 -  93)
   INTEGRAL    Likelihood = -0.69   Transmembrane  229 - 245 ( 229 - 245)

----- Final Results -----
              bacterial membrane --- Certainty=0.7262(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA32095 GB:AB010970 ORF7 [Streptococcus mutans]
 Identities = 374/775 (48%), Positives = 525/775 (67%), Gaps = 7/775 (0%)

Query: 53  VSFVGYIISLIGLSYYLSRQVSRQLFLKTSFIVISYLIVSYWVQITQHLNDKRPDIWSLT 112
           V  V Y++S++GLS+YLS+ + +  F++        Y+++SY++ +T+ LN++ F IW L
Sbjct: 30  VCLVIYVLSILGLSFYLSKNLKKTFFIELLLGYGLYIVISYFLAVTRELNNESFKIWDLA 89

Query: 113 KNQFVQFQALPSLLIILV----MATLIKILAAYFAIRKDRPGLL-GYQGNTFSVALILAV 167
           KN F+Q   LP+L++I+    + LI++       + +   LL  +    F +  ++
Sbjct: 90  KNHFFQPFTLPTLVLIIACTFALNYLIRVKMKRSHLSRKMTLLLSNFSETEFLITGLVVS 149

Query: 168 VPINDIHLLKLISGRFSSLVTWGNSQIALLKISGLLIVLLVIPATIIYVVLNALKHLKSN 227
             ++D   +KL+          +  +LL +  LL   L++F+ I+    NA + +K N
Sbjct: 150 FILSDTLYVKLLQESLRAYYHKPLAYESLLFLYTLIT--LILPSVIVEACFNAYRSIKLN 207

Query: 228 KPSFSVAATTSLFLALVFNYTFQYGVKGDEALLGYYVFPGATLFQIVATTLVALLAYVIT 287
           +P+ S+A  +SL  A +FNY FQYG+K D  LLG Y+ PGAT +QI+ +T      Y+I
Sbjct: 208 RPNLSLAPVSSLLFATIFNYAFQYGLKNDADLLGKYIVPGATAYQLLVITANGFFLYLII 267

Query: 288 NRYWPTTFFLLILGTIISVVNDLKESMRSKPLLVTDFVWLQSLGLVTSFVKKSVIVEMVV 347
           NRY   TF ++ILG+II+VVN LK  MR+KPLLVTDF W+  + L+   V  ++I  ++
Sbjct: 268 NRYLLVTFLIVILGSIITVVNVLKVGMRNSPLLVTDFAWVTNIRLLARSVNANIIFSTLL 327

Query: 348 GLAICIVVAWYLHGRVLAGKLFMSPVKRASAVLGLPIVSCSMLIPFSYEKEGKILSGLPI 407
            LA  I++  +L  R+L GK+  +   +  + ++  S+ I F  EK  KI++G+P+
Sbjct: 328 ILAALILLYLFRKKLLQGKTTENYRLKVGLISSICLLGPSIFIIFRNEKGSKIVNGIPV 387

Query: 408 ISALNNDNDINWLGFSTNARYKSLAYVWTRQVTKKIMSKPTNYSQETIASIAQKYQKLAE 467
           IS +NN  DI + GF +NA YKSL YVWT+QVTK IM+KP++YS+E I  +A+KY  +A
Sbjct: 388 ISQVNNWVDIGYQGFYSNASYKSLMYVWTKQVTKSIMDKPSDYSKERILKLAKKYNNVAN 447
```

```
Query: 468 DINKDRKNNIADQTVIYLLSESLSDPDRVSNVTVSHDVLPNIKAIKNSTTAGLMQSDSYG 527
           INK R  NI++QTVIY+LSES SDPDRV  V +S DV+PNIK IK  TT+GLM SD YG
Sbjct: 448 KINKVRTENISNQTVIYILSESFSDPDRVQGVNLSRDVIPNIKQIKESTSGLMHSDGYG 507

Query: 528 GGTANMEPQTLTSLPFYNFSSSVSVLYSEVPKMAKPHTISEFYQGKNRIAMHPASANNF 587
           GGTANMEPQ+LT LP+YNF+SSVS LY+EV P M+   +IS  ++ KMR+ +HP+SA+N+
Sbjct: 508 GGTANMEPQSLTGLPYYNFNSSVSTLYTEVVPDMSVPFSISNQFKSKNRVVIHPSSASNY 567

Query: 588 NRKTVYSNLGPSKFLALSGSKDKFKNIENVGLLTSDKTVYNNITLSLINDSESQFFSVTTM 647
           +RK VY  L F  F+A SG+ DK  + E VGL  SDKT Y NIL  INPS+SQFFSV+TM
Sbjct: 568 SRKYVYDKLKPTFVASSGTSDKITHSEKVGLNVSDKTTYQNILDKINDSQSQFFSVMTM 627

Query: 648 QNHIPWSSDYPERIVAEGKNFTSENHNLTSYARLLGFTDKETRAFLEKLTQINKPITVV 707
           QNH+PW+SD P ++VA GK +T++EN +L+SYARLL+TDKET+ FL +L+Q+   +TVV
Sbjct: 628 QKHVPWASDEPSDVVATGKGYTKDENGSLSSYARLLTYTDKETKDFLAQLSQLHKVTVV 687

Query: 708 FYGDHLPGLYPDSAFNKHIENKYLTDYFIWSNGTNKKKNEPLINSSDFTAALFEHTDSKV 767
           FYGDHLPGLYP+SAF K  ++Y TDYFIWSN   + NH  +NSSDFTA L EHT+SKV
Sbjct: 688 FYGDHLPGLYPESAFKKDPDSQYQTDYFIWSNYMTKTLNHSYVNSSDFTAELLEHTNSKV 747

Query: 768 SFYYALLTSVLNKASVDKSPDSPEVKAIQNDLKNIQYDVTIGKGYLLKHKTFPKI 822
           SFYYALLTEVL+  +V     + E K I NDLK IQYD+T+GKGY+  +K FF I
Sbjct: 748 SFYYALLTSVLDNTTVGHGKLTKEQKEIANDLKLIQYDITVGKGYIRNYKGFPDI 802
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2615

A DNA sequence (GASx789R) was identified in *S.pyogenes* <SEQ ID 7729> which encodes the amino acid sequence <SEQ ID 7730>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.06   Transmembrane   42 -  58 (  42 -  58)

----- Final Results -----
               bacterial membrane --- Certainty=0.1426(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2616

A DNA sequence (GASx790) was identified in *S.pyogenes* <SEQ ID 7731> which encodes the amino acid sequence <SEQ ID 7732>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2617

A DNA sequence (GASx791) was identified in *S.pyogenes* <SEQ ID 7733> which encodes the amino acid sequence <SEQ ID 7734>. Analysis of this protein sequence reveals the following:

```
Possible site: 48

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-12.42   Transmembrane  166 - 182 ( 157 - 188)
   INTEGRAL    Likelihood = -7.32   Transmembrane   85 - 101 (  79 - 104)
   INTEGRAL    Likelihood = -6.90   Transmembrane  397 - 413 ( 386 - 417)
   INTEGRAL    Likelihood = -6.05   Transmembrane  253 - 269 ( 252 - 273)
   INTEGRAL    Likelihood = -5.26   Transmembrane  301 - 317 ( 293 - 325)
   INTEGRAL    Likelihood = -3.35   Transmembrane  363 - 379 ( 362 - 379)
   INTEGRAL    Likelihood = -3.24   Transmembrane  335 - 351 ( 335 - 351)

----- Final Results -----
               bacterial membrane --- Certainty=0.5967(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA64645 GB:U10927 CapF [Staphylococcus aureus]
 Identities = 97/419 (23%), Positives = 186/419 (44%), Gaps = 40/419 (9%)

Query: 12  FLWNMLGSLSTAVISVILLMVVTRLLTSADSDIYAFAYSFANMMVVGLFQVRNYQATDI 71
           F +  + ++ +A+    ++L+V+ RL T  D    Y +A      + +     ++R+   T
Sbjct: 5   FNYMFVANILSALCKFLILLVIVRLGTPEDVGRYNYALVITAPIFLFISLKIRSVIVT-- 62

Query: 72  NEKYSFSQYLVARLMTCLLMLAITVIYLTLTKTDSYKSTIVFLVCPYRSTDAPSDLYQGM 131
           N+KYS ++Y+ A L    ++ L    I++ +          T + +V   +  +    G+
Sbjct: 63  NDKYSPNEYISAILSLNIITLIFVAIFVYVLGNGDL--TTILIVSLIKLFENIKEVPYGI 120

Query: 132 FQQHERLDIAGKSLAYRNTLIFMVYTAIILYSKNLTLALVAVCIVSLVFIMYYDIGHSKK 191
           +Q++E L + G S+   N L  +++   I  +S NL +AL+ + I  +     D  +  K
Sbjct: 121 YQKNESLKLLGISMGIYNILSLLFYIIYSFSHNLNMALLFLVISCIFSFAIIDRWYLSK 180

Query: 192 PQKLMPSELLSNISPQNSLKLLKESF----PLFLNGFLIIYIYTQPKYAIELMTTLGEVA 247
           +  +          + + N++   KE F    PL  +  L        P+  +E +   G+
Sbjct: 181 YYNI-------KLHYNNNIAKPKEIFILTIPLAPSSALGSLNYGIPRIVLENL--PGKYT 231

Query: 248 LGS-QTIFNILFMPAFVMNLLILFFRPHITQMAIALIRGQIK-EFNKLQVQLFAYLGVF- 304
           LG   TI  +L +      N +   F P + +     L + + K EF K+  ++  ++G+F
Sbjct: 232 LGIFSTIAYVLVIGGLFANSISQVFLPKLRK----LYKDEKKIEFSKLITKM-VFIGIFI 286

Query: 305 SLIALVGSGGLPGIPFLSILYG-----TNLTDYWVDF-MLIMLGGSIGSFATVIDNILTAM 358
            + +++ S   G   LS+L+G      N+    + F +L +L G      +
Sbjct: 287 GMCSVILSLFLGEALLSLLFGKEYGRNNIILLILSPGLLFILSGIFLGTTIIATGKYNVN 346

Query: 359 RKQQLLLIPYTGPLISLLITNLPVMKYHILGAALSFLITMLVWLGLSIMIYLFIMNRF 417
            K  L+L+     F I L+ + L + KY +LGAAL+  I+  V L   I  Y F     F
Sbjct: 347 YKISLILL-----PCI-LIFSFLLIPKYSLLGAALTITISQFVAL---ISYYYFYKRIF 396
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2618

A DNA sequence (GASx792) was identified in *S.pyogenes* <SEQ ID 7735> which encodes the amino acid sequence <SEQ ID 7736>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-10.03   Transmembrane    64 -  80 (  60 -  84)
   INTEGRAL    Likelihood = -9.66   Transmembrane    43 -  59 (  37 -  63)
   INTEGRAL    Likelihood = -8.70   Transmembrane   232 - 248 ( 229 - 251)
   INTEGRAL    Likelihood = -8.28   Transmembrane   410 - 426 ( 402 - 432)
   INTEGRAL    Likelihood = -6.21   Transmembrane   290 - 314 ( 296 - 322)
   INTEGRAL    Likelihood = -6.21   Transmembrane   478 - 494 ( 471 - 496)
   INTEGRAL    Likelihood = -5.04   Transmembrane   265 - 281 ( 256 - 288)
   INTEGRAL    Likelihood = -3.29   Transmembrane   380 - 396 ( 378 - 397)
   INTEGRAL    Likelihood = -2.92   Transmembrane   210 - 226 ( 209 - 227)
   INTEGRAL    Likelihood = -2.60   Transmembrane   187 - 203 ( 187 - 204)
   INTEGRAL    Likelihood = -2.50   Transmembrane   442 - 458 ( 439 - 458)
   INTEGRAL    Likelihood = -1.65   Transmembrane    18 -  34 (  18 -  35)
   INTEGRAL    Likelihood = -1.38   Transmembrane   165 - 181 ( 165 - 181)

----- Final Results -----
               bacterial membrane --- Certainty=0.5012(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAA19642 GB:AB002668 unnamed protein product [Actinobacillus
          actinomycetemcomitans]
 Identities = 116/459 (25%), Positives = 207/459 (44%), Gaps = 60/459 (13%)

Query: 69  FILVFGTISAIISPINDIPDEYVHYSRTVYISEGDINLTNNNKKLRISKDVDKLI----- 123
           FIL F  I  II+P    PDE+ H+ R    IS G I   ++   K   + K + K++
Sbjct: 16  FILTF-IIGVIITPPYQSPDEPYHFQRGYAISNSQIIPSSTEK---LDKVNNNMLSIYEG 71

Query: 124 ----KQSGKTFITSNLKATKHSTREYSYPYIKGTNAYYSPSYTPQALGILVGNALDLPIL 179
                ++  T    N         +EY       TN Y+   Y+PQALG  +G+ LDL +
Sbjct: 72  IPYRSENKVTHFLENEAQNVAWEKEYILDESANTNVYFPLIYLPQALGSPLGSTLDLSLY 131

Query: 180 LTYYPGRLCN-LISYAMLAPIAIKLSGSPKQVIAVVTLLPWNIYLAASFNQDGFAIGLVL 238
             YY  ++   L+S A+L F +++   S    + ++  LPM ++   S N D       ++
Sbjct: 132 NMYYLAKIFTLVVSIAILYPASVQYRLSIP--VLLILSLPMTMFQMGSTNPDS-----II 184

Query: 239 VTIGLPI-NLLSSKDKSNYNTKPFLYLVLCGLL------VLSKFTYFLLVCLPLPIPNEK 291
            ++ +PI +LL+    SNYN    F +    C LL      V  KF   +L+ LP FI   +
Sbjct: 185 FSLSVFIGSLLARGLDSNYN---PTHKDPCKLLFSIFLCVTVKFNMMLVLLLLPFFISKRR 241

Query: 292 PGKNTKLVILKKLGGLLIFLFAAMWFRLYGQVKTPVVADPLKEV----MVSQQVKNMLE 347
            ++  +  +  +L  + A          K   + +F   +     ++ + KN L
Sbjct: 242 EIRHGSMYSIFIIILSILWIVLAMKLTEAQSHFKRGAIHNPSYYYIPHMDDLPEIPKNTLN 301

Query: 348 SPIVYSSIIIRHMVINLINMNNIFQFGA-LSYGITNLFPLYVCFFFVYISNASKITTINI 406
            + Y   ++R  +  L  ++  F     L +G T+L       + F++I N  K+   I
Sbjct: 302 --LYLKSLLRMPLGVLGWVDTKFTINEYLPFGSTSLLA-----YIPLFTHNLYKLKYVI 354

Query: 407 VEKM--GIIFVISAIIGATVLAMYLTWTPVGSSTVLGVQSRYLIGIIPLVLLLFSS---- 460
             V  +  G++F+ +  I     + +T+  +G++ ++GVQ RY    IP++L++FSS
Sbjct: 355 VSVLLVGVVFLFTHFI------LLTTYNEIGTTQIVGVQGRY---FIPIMLIIFSSFILK 405
```

```
Query: 461 QQQKFKQIEDILSDKLAIHVSLLFILAMLM--STIFRYY 497
           + +K    + I    + +    LFI + +   + + RYY
Sbjct: 406 KSEKTSNNKTISKYFIIVPFLFLFISSFTITNTLVSRYY 444
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2619

A DNA sequence (GASx797) was identified in *S.pyogenes* <SEQ ID 7737> which encodes the amino acid sequence <SEQ ID 7738>. Analysis of this protein sequence reveals the following:

```
Possible site: 49

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1491(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC83961 GB:L47648 cytidine monophosphate kinase [Bacillus subtilis]
 Identities = 116/220 (52%), Positives = 156/220 (70%), Gaps = 1/220 (0%)

Query: 2   KAIKIAIDGPASSGKSTVAKIIAKNLGYTYLDTGAMYRSATYIALTHGYTGKEVALILEE 61
           K + IAIDGPA++GKSTVAKI+A+   Y Y+DTGAMYR+ TY AL       +   + E
Sbjct: 3   KKLSIAIDGPAAAGKSTVAKIVAEKKSYIYIDTGAMYRAITYAALQENVDLTDEEKLAEL 62

Query: 62  LEKNPIFFKKAKDGSQLVFLGDEDVTLAIRQNDVTNNVSWISALPEIREKLVHQQRRIAQ 121
           L++  I    KDG Q VF+   DVT AIR +++N VS  +    +REK+V +Q++ + +
Sbjct: 63  LKRTDIELITTKDG-QKVFVNGTDVTEAIRTDSISNQVSIAAKHRSVREENVERQQQLGE 121

Query: 122 AGGIIMDGRDIGTVVLPDAELKIFLVASVEERAERRYKENLEKGIESDFETLKEEIAARD 181
            GG++MDGRDIGT VLP+AE+KIFL+ASVEERA+RRY+EN++KG + ++ETL EEIA RD
Sbjct: 122 KGGVVMDGRDIGTHVLPNAEVKIFLLASVEERAKRRYEENVKKGFDVNYETLIEEIARRD 181

Query: 182 YKDSHRKVSPLKAAEDALIPDTTGVSIDGVVQFIQEKAEK 221
            DS R+VSPL+ AEDAL  DTT +SI  V   I E  E+
Sbjct: 182 KLDSEREVSPLRKAEDALEIDTTSLSIQEVADKILEAVEQ 221
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2620

A DNA sequence (GASx799) was identified in *S.pyogenes* <SEQ ID 7739> which encodes the amino acid sequence <SEQ ID 7740>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4324(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA34313 GB:X16188 ribosomal protein L35 (AA 1-66) [Bacillus
          stearothermophilus]
 Identities = 46/65 (70%), Positives = 51/65 (77%)

Query: 1  MPKQKTHRASAKRFKKTGSGGLKRFRAFTSHRFIGKTKKQRRHLRKAGLVSGGDFKRIKA 60
          MPK KTHR SAKRFK+T SG LKR  A+TSH F  KTKKQ+RHLRKA LVS GDFKRI+
Sbjct: 1  MPKMKTHRGSAKRFKKTASGKLKRGHAYTSHLFANKTKKQKRHLRKATLVSPGDFKRIRQ 60

Query: 61 MVTGL 65
          M+  L
Sbjct: 61 MLDNL 65
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2621

A DNA sequence (GASx806R) was identified in *S.pyogenes* <SEQ ID 7741> which encodes the amino acid sequence <SEQ ID 7742>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.5361(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2622

A DNA sequence (GASx809R) was identified in *S.pyogenes* <SEQ ID 7743> which encodes the amino acid sequence <SEQ ID 7744>. Analysis of this protein sequence reveals the following:

```
Possible site: 52

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -8.81   Transmembrane   33 -  49 (  28 -  53)

----- Final Results -----
              bacterial membrane --- Certainty=0.4524(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2623

A DNA sequence (GASx814R) was identified in *S.pyogenes* <SEQ ID 7745> which encodes the amino acid sequence <SEQ ID 7746>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0206(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2624

A DNA sequence (GASx817) was identified in *S.pyogenes* <SEQ ID 7747> which encodes the amino acid sequence <SEQ ID 7748>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -1.49   Transmembrane   16 -  32 (  15 -  32)

----- Final Results -----
            bacterial membrane --- Certainty=0.1595(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2625

A DNA sequence (GASx820) was identified in *S.pyogenes* <SEQ ID 7749> which encodes the amino acid sequence <SEQ ID 7750>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -7.11   Transmembrane   62 -  78 (  59 -  81)
   INTEGRAL    Likelihood = -6.00   Transmembrane  128 - 144 ( 123 - 147)
   INTEGRAL    Likelihood = -2.50   Transmembrane    5 -  21 (   3 -  26)

----- Final Results -----
            bacterial membrane --- Certainty=0.3845(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA26653 GB:M83994 prolipoprotein signal peptidase
          [Staphylococcus aureus]
 Identities = 57/153 (37%), Positives = 96/153 (62%), Gaps = 6/153 (3%)

Query: 1   MKKRLFVLSLLLL----VALDQLSKFWIVSHIALGEVKDFIPGIVSLTYLQNNGAAFSIL 56
            M K+ F+ + IL+    V  DQ++K+ I + + +G+    IP  ++++T  +NNGAA+ IL
Sbjct: 1   MHKKYFIGTSILIAVFVVIFDQVTKYIATTMKIGDSFEVIPHFLNITSHRNNGAAWGIL 60

Query: 57  QDQQWFFVVITVLVIGYAIYYLATHPHLNIWKQLALLLIISGGIGNFIDRLRLAYVIDMI 116
             +  FF +IT++++   +Y+       N++ Q+A+ L+ +G +GNFIDR+    V+D I
Sbjct: 61  SGKMTFFFIITIILLALVYFFIKDAQYNLPMQVAISLLFAGALGNFIDRILTGEVVDFI 120

Query: 117 HLDF--VDFAIFNVADSYLTVGVILLLICLWKE 147
             +    DF IFN+ADS LT+GVIL++I L K+
Sbjct: 121 DTNIFGYDFPIFNIADSSLTIGVILIIIALLKD 153
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2626

A DNA sequence (GASx822R) was identified in *S.pyogenes* <SEQ ID 7751> which encodes the amino acid sequence <SEQ ID 7752>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2638(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2627

A DNA sequence (GASx823R) was identified in *S.pyogenes* <SEQ ID 7753> which encodes the amino acid sequence <SEQ ID 7754>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3452(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2628

A DNA sequence (GASx828) was identified in *S.pyogenes* <SEQ ID 7755> which encodes the amino acid sequence <SEQ ID 7756>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2629

A DNA sequence (GASx836) was identified in *S.pyogenes* <SEQ ID 7757> which encodes the amino acid sequence <SEQ ID 7758>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial cytoplasm --- Certainty=0.4333(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2630

A DNA sequence (GASx853R) was identified in *S.pyogenes* <SEQ ID 7759> which encodes the amino acid sequence <SEQ ID 7760>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial cytoplasm --- Certainty=0.4906(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2631

A DNA sequence (GASx854R) was identified in *S.pyogenes* <SEQ ID 7761> which encodes the amino acid sequence <SEQ ID 7762>. Analysis of this protein sequence reveals the following:

```
Possible site: 43

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3989(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

A related sequence was also identified in GAS <SEQ ID 9147> which encodes the amino acid sequence <SEQ ID 9148>. Analysis of this protein sequence reveals the following:

```
Possible site: 42

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty= 0.399(Affirmative) < succ>
               bacterial membrane --- Certainty= 0.000(Not Clear) < succ>
                bacterial outside --- Certainty= 0.000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB59092 GB:M97157 pyrogenic exotoxin C [Streptococcus pyogenes]
 Identities = 39/67 (58%), Positives = 53/67 (78%)

Query: 1   LMESKEIYLTKSPYIRGSLEIHSKNRKHEKINLYDAKPNSTRSDVFKKYKDNKTINMKDF 60
           LM++ +IY    SPY+ G +EI +K+ KHE+I+L+D+     TRSD+F KYKDN+ INMK+F
Sbjct: 167 LMDNYKIYDATSPYVSGRIEIGTKDGKHEQIDLFDSPNEGTRSDIFAKYKDNRIINMKNF 226

Query: 61  SHFDIYL 67
           SHFDIYL
Sbjct: 227 SHFDIYL 233
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2632

A DNA sequence (GASx855R) was identified in *S.pyogenes* <SEQ ID 7763> which encodes the amino acid sequence <SEQ ID 7764>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2633

A DNA sequence (GASx856) was identified in *S.pyogenes* <SEQ ID 7765> which encodes the amino acid sequence <SEQ ID 7766>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4145(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2634

A DNA sequence (GASx862) was identified in *S.pyogenes* <SEQ ID 7767> which encodes the amino acid sequence <SEQ ID 7768>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.6285(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2635

A DNA sequence (GASx863) was identified in *S.pyogenes* <SEQ ID 7769> which encodes the amino acid sequence <SEQ ID 7770>. Analysis of this protein sequence reveals the following:

```
Possible site: 51

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2636

A DNA sequence (GASx878) was identified in *S.pyogenes* <SEQ ID 7771> which encodes the amino acid sequence <SEQ ID 7772>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2637

A DNA sequence (GASx887R) was identified in *S.pyogenes* <SEQ ID 7773> which encodes the amino acid sequence <SEQ ID 7774>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1911(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2638

A DNA sequence (GASx910) was identified in *S.pyogenes* <SEQ ID 7775> which encodes the amino acid sequence <SEQ ID 7776>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4511(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2639

A DNA sequence (GASx911) was identified in *S.pyogenes* <SEQ ID 7777> which encodes the amino acid sequence <SEQ ID 7778>. Analysis of this protein sequence reveals the following:

```
Possible site: 52

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2993(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC74707 GB:AE000259 glutathionine S-transferase [Escherichia
          coli]
 Identities = 29/137 (21%), Positives = 61/137 (44%), Gaps = 9/137 (6%)

Query: 1   LPFIAKQTLKSQLIPQDNLLAESRFNEIMDFLTGDFFLVFRPMINPHRYTISQDNQALEK 60
           + ++A      QL+   N ++  +  E ++++  +    F P+  P            E+
Sbjct: 70  MQYLADSVPDRQLLAPVNSISRYKTIEWLNYIATELHKGFTPLFRP---------DTPEE 120

Query: 61  VKQASYKRMDIAMTHLDSLIGESGHVYVRDQQTIADAYAYAMALWSQKTPKSYENYPHLAA 120
           K     +++  + +++  + +   +  + TIADAY + +  W+     + E   H+AA
Sbjct: 121 YKPTVRAQLEKKLQYVNEALKDSHWICGQRPTIADAYLFTVLRWAYAVKLNLEGLEHIAA 180

Query: 121 FMAKMVEDSAVQQVLNA 137
           FM +M E   VQ  L+A
Sbjct: 181 FMQRMAERPEVQDALSA 197
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2640

A DNA sequence (GASx932R) was identified in *S.pyogenes* <SEQ ID 7779> which encodes the amino acid sequence <SEQ ID 7780>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4081(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2641

A DNA sequence (GASx935) was identified in *S.pyogenes* <SEQ ID 7781> which encodes the amino acid sequence <SEQ ID 7782>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.6304(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2642

A DNA sequence (GASx937) was identified in *S.pyogenes* <SEQ ID 7783> which encodes the amino acid sequence <SEQ ID 7784>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3503(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2643

A DNA sequence (GASx938R) was identified in *S.pyogenes* <SEQ ID 7785> which encodes the amino acid sequence <SEQ ID 7786>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2884(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2644

A DNA sequence (GASx939) was identified in *S.pyogenes* <SEQ ID 7787> which encodes the amino acid sequence <SEQ ID 7788>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2771(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2645

A DNA sequence (GASx941) was identified in *S.pyogenes* <SEQ ID 7789> which encodes the amino acid sequence <SEQ ID 7790>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2257(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2646

A DNA sequence (GASx942R) was identified in *S.pyogenes* <SEQ ID 7791> which encodes the amino acid sequence <SEQ ID 7792>. Analysis of this protein sequence reveals the following:

```
Possible site: 23

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3255(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB91582 GB:AF242881 vmh [Agrobacterium tumefaciens] (ver 2)
 Identities = 75/223 (33%), Positives = 116/223 (51%), Gaps = 2/223 (0%)

Query: 38  DQNSGFNKHKEVVNLVSDILNRTQNTDNIKLVIEYVCNPLRYINEVSIFEQLRTAINIPL 97
            D +    K   R++N  +   N  +   +I   I     P R+  +   FE +R  +N  L
Sbjct: 39  DTDPQMTKHRELYNAFASDQNSRKQRTHIIAFIRKAMKPERFARDSEKFEPMRLNLNRAL 98

Query: 98  SLKGLIVSDSGQIVTTTSKTLSEAKKRFETLDSRLKELKVHPHVLKFCTQELLQENYFH 157
            +  GL V  SG++     ++TLS+A +R   L + L     VHP VL+FC +ELL +NYFH
Sbjct: 99  APAGLAVKASGELAAVDAAETLSQATRRALELRADLTSRGVHPDVLRFCREELLVDNYFH 158

Query: 158 AVPEASKGVFHRIRLLTGSAMDSASLIDQCFKPGEPIVIINGNKLQTLDEQSEYKGLKNL 217
            AV EA K V  +IR  TG   D A L+D+ F    P++ I  N+LQ+  E+ E +G  NL
Sbjct: 159 AVLEAVKSVADKIRQRTGLTDDGAVLVDRAFSGDAPMLAI--NELQSESEKGSQRGFSNL 216

Query: 218 LLAIAHLYRNSKAHKLKYYNPDNLNDALTALTLMSLAHNLLDS 260
            +    ++EN+ AH  + +   +  DA    ++ SL H  +D+
Sbjct: 217 VKGTFSMFRNTTAHAPRIHWQMSKEDAEDLFSMFSLMHRRIDA 259
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2647

A DNA sequence (GASx943R) was identified in *S.pyogenes* <SEQ ID 7793> which encodes the amino acid sequence <SEQ ID 7794>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1526(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2648

A DNA sequence (GASx944) was identified in *S.pyogenes* <SEQ ID 7795> which encodes the amino acid sequence <SEQ ID 7796>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1427(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2649

A DNA sequence (GASx945) was identified in *S.pyogenes* <SEQ ID 7797> which encodes the amino acid sequence <SEQ ID 7798>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2578(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC98430 GB:L29324 excisionase [Streptococcus pneumoniae]
 Identities = 23/54 (42%), Positives = 40/54 (73%)

Query: 1  LIQQWEGLTVATAKQWATEMRDHPDFKQFVLNPTHRIVFIDYEGPKLFVQWKSR 54
          ++++W+GL   T  +W  EMR++  F  +V+NPTH++VFI+ EGF+ F++WK +
Sbjct: 21 ILKRWDGLNKYTLNRWIKEMRENKTFSMYVINPTHKLVFINLEGFSSFLRWKQK 74
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2650

A DNA sequence (GASx946) was identified in *S.pyogenes* <SEQ ID 7799> which encodes the amino acid sequence <SEQ ID 7800>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -4.99   Transmembrane    3 -  19 (   1 -  23)

----- Final Results -----
             bacterial membrane --- Certainty=0.2996(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2651

A DNA sequence (GASx950) was identified in *S.pyogenes* <SEQ ID 7801> which encodes the amino acid sequence <SEQ ID 7802>. Analysis of this protein sequence reveals the following:

```
Possible site: 51

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2211(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2652

A DNA sequence (GASx951) was identified in *S.pyogenes* <SEQ ID 7803> which encodes the amino acid sequence <SEQ ID 7804>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4258(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2653

A DNA sequence (GASx952) was identified in *S.pyogenes* <SEQ ID 7805> which encodes the amino acid sequence <SEQ ID 7806>. Analysis of this protein sequence reveals the following:

```
Possible site: 46

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2476(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF74110 GB:AF212847 ORF245 [Lactococcus lactis bacteriophage
            ul36.2]
 Identities = 82/265 (30%), Positives = 128/265 (47%), Gaps = 27/265 (10%)

Query: 1   MANQLSTQQVKRDITTDPTLLTGADIKKYFDPQNLLSKKQVQQALALCKGRNLNPFANEV 60
           MAN+L    V          L    IK+Y D     S+ ++   + LCK  N+NPF  EV
Sbjct: 1   MANELGIFSVDN--------LNMTTIKQYLDGGGKASDAELVLLINLCKQNNMNPFMKEV 52
```

```
Query: 61  YIVAYKNNSGTDFSLIVSKRAFMKRAERCEGYDGFEAGITVM-RNGEMVEIRGSLKLPDD 119
           Y + Y N          ++VS++ + KRA +   + G E G+ V+ ++G +   EG+ K  +
Sbjct: 53  YFIKYGNQPA---QIVVSKDFYRKRAFQNPNFVGIEVGVIVLNKDGVLEHNEGTFKTHEQ 109

Query: 120 VLIGGNAIVYRKDRSHRYKVTVDPNEYVKLDKYGNPRSTWKSMPGTMIRKTALVQTLREA 179
            L+G WA V+ K+       V V ++EYV++ K G+P   W + P TM+ K A  Q LR A
Sbjct: 110 ELVGAWARVHLKNTEIPVYVAVSYDEYVQM-KDGHPNKMWTNKPCTMLGKVAESQALRMA 168

Query: 180 FPDELGNMYTDIDGGDTFDAIKDVTPQETQEEVRARK---MAQIEQYKQEQ--TQKQTQK 234
           FP E    Y + +  +          P++   EV   K    AQIE + +E    +K  +
Sbjct: 169 FPAEFSGTYGKEEYPE---------PEKEPREVNGVKESPDRAQIESFDKEDYAAKKIEEL 219

Query: 235 ADTSYPVDEVSEHTDDPVQGKLLDG 259
            + + P  EV E T + +  E L+G
Sbjct: 220 KEKAQPQKEVVEETGEVIDEEPLEG 244
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2654

A DNA sequence (GASx953) was identified in *S.pyogenes* <SEQ ID 7807> which encodes the amino acid sequence <SEQ ID 7808>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3413(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF74111 GB:AF212847 ORF364 [Lactococcus lactis bacteriophage
          ul36.2]
 Identities = 67/222 (30%), Positives = 120/222 (53%), Gaps = 3/222 (1%)

Query: 1   MQELQLKVTQAQVEIIDREKTEQNINEVVAKYQNYAVTAGTIKDDKQVLADLRKLKKQLS 60
           ++++++     A + I++ EKP+ +IN+VVA+Y  +  +  +  D++  A L KL  ++
Sbjct: 19  VKDIEIDFXPAIINILEESKFKASINQVVAEYTGHVPSVSNLYVDRKTRASLNKLIYKIE 78

Query: 61  DERIKVKKSLSKPADDIDGYIKQASKPLDDTIDKIATDVKRFEDHQKALRLDTVKSYLSN 120
            R ++KK ++ P  + +G+ K+A  P++  I+ I   +K+ E  QK  R   V   L
Sbjct: 79  TRRKEIKKSINVPYAEPEGWYKKAIAPMEKVIETIDAGIKKIEAEQKESRKKVVHELLVE 138

Query: 121 KASEYMLDPRIFDRKAMEYTKAGNFMADGVTLKKVTMKSLEDLVTPEYQKEQEVEKAKAT 180
             ++  +D RIP+     ++ K+ NF  + +  KK  + S+  ++  E QK  E + AK +
Sbjct: 139 LTTDTEVDSRIFENFVDDWAKSSNF--NDIKPKKQLIDSITYVIDGEKQKIAEYKSAKQS 196

Query: 181 ISGQCAEYGMTDQPYIRMLKE-MTLVEVLGQIKADYLAEKQK 221
           IS  C   +T  PYIRML    T+ E++  I  D L EKQ+
Sbjct: 197 ISDFCFGNNITSTPYIRMLDSGKTVSEIMAVITEDVLFEKQR 238
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2655

A DNA sequence (GASx954) was identified in *S.pyogenes* <SEQ ID 7809> which encodes the amino acid sequence <SEQ ID 7810>. Analysis of this protein sequence reveals the following:

```
Possible site: 56

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3884(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2656

A DNA sequence (GASx955) was identified in *S.pyogenes* <SEQ ID 7811> which encodes the amino acid sequence <SEQ ID 7812>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1777(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2657

A DNA sequence (GASx956) was identified in *S.pyogenes* <SEQ ID 7813> which encodes the amino acid sequence <SEQ ID 7814>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -2.44   Transmembrane    82 -  98 (  81 -  98)

----- Final Results -----
             bacterial membrane --- Certainty=0.1977(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2658

A DNA sequence (GASx958) was identified in *S.pyogenes* <SEQ ID 7815> which encodes the amino acid sequence <SEQ ID 7816>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3673(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2659

A DNA sequence (GASx960) was identified in *S.pyogenes* <SEQ ID 7817> which encodes the amino acid sequence <SEQ ID 7818>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1852(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2660

A DNA sequence (GASx961) was identified in *S.pyogenes* <SEQ ID 7819> which encodes the amino acid sequence <SEQ ID 7820>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.7380(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF63071 GB:AF158600 gp137 [Streptococcus thermophilus
          bacteriophage Sfi11]
 Identities = 57/136 (49%), Positives = 97/136 (71%), Gaps = 2/136 (1%)

Query: 5   PEIDIQKTKSNAKRKLREYPRWRRIANDVDTQKVTATYSPERDQPHGTPSKPVERLALNR 64
           PEID + T     KRKLREYPRWR IA+D   QK+T  ++P PR   G  +KPVE +A+ R
Sbjct: 4   PEIDEKATLKRCKRKLREYPRWRETAHDSABQKITQEPTFMPRG--GGVNKPVENIAVRR 61

Query: 65  VSAEQELDTIERAVNGIFDPEYRLILIDKYLLTYPKTDCDIYTKLGYEKSQYYNMLDNAL 124
           V A  EL+ IE+AVNG++ P+YR ILI+KYL   PK +  I   +G+E++ +  +L+N++
Sbjct: 62  VDALNELEAIEQAVNGLYKPDYRRILIEKYLAYPPKPNWQIAQSIGFERTAFQELLNNSI 121

Query: 125 LSFSELYKEGMLLVEK 140
           L+F+ELY++G L+VE+
Sbjct: 122 LAFAELYRDGRLIVER 137
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2661

A DNA sequence (GASx962) was identified in *S.pyogenes* <SEQ ID 7821> which encodes the amino acid sequence <SEQ ID 7822>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3375(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2662

A DNA sequence (GASx963R) was identified in *S.pyogenes* <SEQ ID 7823> which encodes the amino acid sequence <SEQ ID 7824>. Analysis of this protein sequence reveals the following:

```
Possible site: 48

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2663

A DNA sequence (GASx964) was identified in *S.pyogenes* <SEQ ID 7825> which encodes the amino acid sequence <SEQ ID 7826>. Analysis of this protein sequence reveals the following:

```
Possible site: 51

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -6.16   Transmembrane   90 - 106 (  89 - 111)
   INTEGRAL    Likelihood = -5.52   Transmembrane  131 - 147 ( 129 - 150)
   INTEGRAL    Likelihood = -0.43   Transmembrane   53 -  69 (  52 -  69)

----- Final Results -----
              bacterial membrane --- Certainty=0.3463(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2664

A DNA sequence (GASx965) was identified in *S.pyogenes* <SEQ ID 7827> which encodes the amino acid sequence <SEQ ID 7828>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3944(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA56779 GB:X98106 Rorf172 [Bacteriophage phig1e]
 Identities = 36/82 (43%), Positives = 52/82 (62%), Gaps = 3/82 (3%)

Query: 18 ELTSKQQRFVDKYITTFNATESAKQAGYSEKSAYSQCQRLLKNVEIQKAMKERFLEAKDT 77
          +LT KQC+F D+YI + NA +++A++AGYS++SA S GQ  L   +I++ + ER     +
Sbjct: 4  KLTPKQCKFADEYIKSGNAADAARKAGYSKRSARSVGQENLTKPDIKQYIDERM---DSI 60

Query: 78 KGDRIQDVAETLEQDTSIARGE 99
              RI D  E +E  T IARGE
Sbjct: 61 ASKRIMDATEAVELLTRIARGE 82
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2665

A DNA sequence (GASx966) was identified in *S.pyogenes* <SEQ ID 7829> which encodes the amino acid sequence <SEQ ID 7830>. Analysis of this protein sequence reveals the following:

```
Possible site: 36
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2389(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB13115 GB:Z99110 PBSX defective prophage terminase (large
            subunit) [Bacillus subtilis]
 Identities = 117/417 (28%), Positives = 195/417 (46%), Gaps = 33/417 (7%)

Query: 31  YRVVKGSERGSKKSKTTALNFIVRLLKYLPWANLLVIRRYSNTNKQSTYTDFKWACNQLKVT 90
           Y+ + G  GS KS   TAL  +++LLK       LVIR    +T++ ST+  F+     +L +T
Sbjct: 21  YQFLVGGYGSSKSYHTALKIVLKLLKEK-RTALVIREVFDTHRDSTFALFQEVIEELGLT 79

Query: 91  HLFKFNESLPEITVKATGQKILFRGLDDELKITSITVDVGALCWAWFEEAVQIETEDKFS 150
               S  ++     G +I+F+G+D+  K+ S    V  +   W EE  +++ E
Sbjct: 80  KAVASLSSPLQLRFH-NGSRIMFKGMDNPAKLKS----VHNISLIWIEECSEVKYEQ--- 131

Query: 151 TVVESIRGSLDADPDFKQITVTFNPWSERHWLKRVFFDEETKR---------------- 193
               + + G L  P+     +  T NP    +W  R FF +E K+
Sbjct: 132 --FKELIGRLRHPKLKLHMICTTNPVGTSNWTYRHFRDERKKRFVLDDSKLYERRTIVK 189

Query: 194 ADTFSGTTTFRVNEWLDDVDKRRYEDLYKTNPRRARIVCDGEWGVAEGLVFDNPEVVDFD 253
            DT+   +T   N +L +   ++ + L + +P   RI   G +GV    V   FEV+  D
Sbjct: 190 GDTYYHHSTANDNLFLPESYVKQLDGLKEYDPDLYRIARGRFGVNGIRVLPQFEVLPHD 249

Query: 254 -VEKTIQRVKET--SAGMDFGFTQDPTTLICVAVDLANKELWLYNEHYQKAMLTDHIVKM 310
             V+K I  + +     GMDFGF +    ++ +AVD   K L++Y E+YQ  M  D  +
Sbjct: 250 QVKKCIAAISKPIFRTGMDFGFEESYNAVVRLAVDPEKKYLYIYWEYYQNKMTDDRTAEE 309

Query: 311 IRDKNLHRSYIAGDSAEKRLIAKIKSKGVSGIVPSIKGKGSIMQGIQPMQGFF-KIYIHPS 369
               +R+    + I  DSAE + I    + +G   +V + K  GS +Q  + ++ F KI+
Sbjct: 310 LREFIETQELIKADSAEPKSIQYFRQQGFR-MVGARKFPGSRLQYTKKVRFKKIPCSDR 368

Query: 370 CEHTIEEFNTYTFKQDKRGNWLNEPIDKNNHVIDAIRYALEKYHIRSNESNQPEVLR 426
               CE+ I E  T T+ +DK G  + +    + H + AI YAL+ Y +    +  + +R
Sbjct: 369 CENVIYELETLTYAKDKNGALIEDEFTIDPHTLSAIWYALDDYEVADMKETAHKRMR 425
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2666

A DNA sequence (GASx967) was identified in *S.pyogenes* <SEQ ID 7831> which encodes the amino acid sequence <SEQ ID 7832>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.4899(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC34397 GB:AF158600 gp502 [Streptococcus thermophilus
```

```
            bacteriophage Sfi11]
 Identities = 67/114 (58%), Positives = 83/114 (72%)

Query: 6   FRDSTGKTKTLEFRFHREARMRYQABSLESLLTEKYLLRREMIEHHDKVQKPRIQELLDY 65
           F DSTG+    L  RFHRE+R+RY+A++LE L+    ++LL+  I HH    Q PRIQELLDY
Sbjct: 7   FTDSTGQDLVLNLRFHRESRIRYRADNLEELMVNNWELLKNFINHHKLRQAPRIQELLDY 66

Query: 66  AEGNNHTISEIGRRKDDDNADVRAVHNYGKYISTLKQQYLVGNPIRVEYIDGTE 119
           A G NH + + GRRKD++NAD RAVHNYG+ IS  K GYL GNPIRVEY D  +
Sbjct: 67  ARGENHDVLKSGRRKDNENADKRAVHNYGRMISKFKTGYLAGNPIRVEYDDNED 120
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2667

A DNA sequence (GASx968) was identified in *S.pyogenes* <SEQ ID 7833> which encodes the amino acid sequence <SEQ ID 7834>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4007(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC34397 GB:AF158600 gp502 [Streptococcus thermophilus
            bacteriophage Sfi11]
 Identities = 172/319 (53%), Positives = 227/319 (70%), Gaps = 9/319 (2%)

Query: 1   LIYRSMDDKTEVVRLDPRSVFVIYQNNLEQSSLAGVRYYNKNQLDGTTKIVELYTDNKIL 60
           +IYRS  D+T + RL P E FVIY N+LE +S+A VRYYN+  L      +VE+YT+  I
Sbjct: 157 VIYRSEYDETRIKRLSPLETFVIYDNSLEDNSIAAVRYYNRGTLQNAKDVVEIYTNQHIY 216

Query: 61  KFEYDGDLTPIGETSSHAFGSVPITEYLNTDDGWGDYETELSLIDLYDAAQSDTANYMQD 120
             +       I  T  HAFG+VPITE+LN   DS+GDYETEL LIDLYD+A+SDTAN+M D
Sbjct: 217 TLDASDSFNEISVTP-HAFGTVPITEFLNNADSIGDYETELYLIDLYDSAESDTANHMSD 275

Query: 121 LSDAILAIIGRVSFPGYVDTAEKAIEYLRKMRKARLLNLEPPVDQDGREGSVDAKYLYKQ 180
           ++DAILAI G ++ P  +  ++        M++ RL+ L+PP   DG+EG+V A+YL K
Sbjct: 276 MADAILAIYGDLALPQGMQASD--------MKRTKLMQLKPPKSADGKEGTVKAEYLTKS 327

Query: 181 YDVQGTEAYKNRIVSDIHKFTNTPDMTDSKPAGQQSGKALKWKVPGLDQERVDMQALFEQ 240
           YDV G EAYK R+  DIH FTNTPDM+D+ P+G  SGKALK+K+PGLDQ+RVD Q+ F Q
Sbjct: 328 YDVSGAEAYKTRLNKDIHVFTNTPDMSDNHPSGNASGKALKYKLPGLDQDRVDTQSQFTQ 387

Query: 241 SLKRRYKLIARVSQLLKEIDDPDISKLKITTPTPNLPKSLQEKIEAFKALGGKLSQETAMA 300
            LKRRY+L AR+  L+ E  DPD S+LKITTPTPNLPKSL E++      LGG++SQETA++
Sbjct: 388 GLKRRYRLAARIGSLVNEFKDPDESRLKITTPTPNLPKSLYEQVSLINDLGGQVSQETALS 447

Query: 301 ITDIVEDAKKEISLINSES 319
           ++ +VE+  +E+  IN ES
Sbjct: 448 LSGLVENPTEELDKINEES 466
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2668

A DNA sequence (GASx969) was identified in *S.pyogenes* <SEQ ID 7835> which encodes the amino acid sequence <SEQ ID 7836>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.5307(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC79543 GB:U68974 ORF28 [Streptococcus thermophilus temperate
          bacteriophage O1205]
 Identities = 118/309 (38%), Positives = 183/309 (59%), Gaps = 18/309 (5%)

Query: 8   YWRDRIKKEMDAK-EADDISLEQSMKQLHDYHFRNIEKEIESFYQRYADKEKIDLSEARK 66
           YW  R  +E +A  +  +    ++ ++ L++      + KR++++ Q+YA+K  + +S+A++
Sbjct: 3   YWSKRTLRERERASIKKGEABFKKELEALYNLQLSQLRKELDAYIQKYANKNGLSVSDAKR 62

Query: 67  RASELDISAYQKKAKELVAKAEKLRREGKIVTRDDFTHQENADMSIYNLAMKTNALELLR 126
           +A   D+ A++ KAK  VA                 DF+ + N ++  YN +M      ELL
Sbjct: 63  KADSFDVKAFETKAKRYVADK------------DFSPKANRELQDYNFSMSVGRQELLI 109

Query: 127 LNIDLEMQRLANGEHHKLTKKFLDEGYRKETEPQAGLLGLSVASQASVKSLADAVINANPK 186
             ++LE+  L+  E +LT  +L  GY+ E   +  LL  +V S  +++     A +NANP+
Sbjct: 110 QELELRLLALSRSERQLTNDYLTNGYKSEV-VRESLLDQTVPSGKTLEKVMKAAVNANPE 168

Query: 187 GAKWSDNIWDRQDKLRSIISQSVQGAILKGKMGLTIARDIRREFDVSASYAKRLAITEHA 246
           GA+WS+ IW RQ++LR I+   V  A+++G+NGLTIAR IR+  D S + A+RLAITEHA
Sbjct: 169 GAEWSERIWKRQEQLRKIVKTEVTRALIRGENGLTIARRIRKHMDASRTEARRLAITEHA 228

Query: 247 RVQMEVGRLSMAENGFAMPDILPEPKACDVCKDIAKH---GPYHLDKWRIGENSPPHPY 303
           RVQ       M ENGF  F ++PE +ACD+CKDI K     P  +     IG N+PP HPY
Sbjct: 229 RVQTLAQESIMKENGFEHPKLMPESRACDICKDIGKETEKNPVKIADMEIGTNAPPIHPY 288

Query: 304 CRCAIVGVD 312
           CRCA+V V+
Sbjct: 289 CRCAVVEVE 297
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2669

A DNA sequence (GASx970) was identified in *S.pyogenes* <SEQ ID 7837> which encodes the amino acid sequence <SEQ ID 7838>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2091(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2670

A DNA sequence (GASx971) was identified in *S.pyogenes* <SEQ ID 7839> which encodes the amino acid sequence <SEQ ID 7840>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
              bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
           bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2671

A DNA sequence (GASx972) was identified in *S.pyogenes* <SEQ ID 7841> which encodes the amino acid sequence <SEQ ID 7842>. Analysis of this protein sequence reveals the following:

```
Possible site: 46

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
           bacterial cytoplasm --- Certainty=0.3226(Affirmative) < succ>
            bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2672

A DNA sequence (GASx973) was identified in *S.pyogenes* <SEQ ID 7843> which encodes the amino acid sequence <SEQ ID 7844>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
           bacterial cytoplasm --- Certainty=0.1830(Affirmative) < succ>
            bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2673

A DNA sequence (GASx975) was identified in *S.pyogenes* <SEQ ID 7845> which encodes the amino acid sequence <SEQ ID 7846>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4757(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB07248 GB:AP001519 unknown [Bacillus halodurans]
 Identities = 46/134 (34%), Positives = 73/134 (54%)

Query: 23  KQPQDEKKYTDADVDAIIDKKFAKWKSEQEAEKSEAKKMAKMNEKEKADYEKQKLLDELQ 82
           K  + E+ +T  +V+ I+  + A+   ++E    EA+K+AKMN ++K +YE +KL  E +
Sbjct: 66  KPNKTERLFTQEEVNRIVKDRLARALKDKEEAIKEAEKLAKMNAEQKREYELEKLRRENS 125

Query: 83  ELKNDKTRNELTAVARQMFAESEINVNDDVLGLVVTLDAEQTKNAVTTLANAFAKVLADD 142
           +LK  + R EL   A +M  E+ I  +DDVL  VV  DAEQT+  V T  +   K+
Sbjct: 126 QLKKAQMKYSLGREATKMLGEAGIMADDDVLSFVVRDDAEQTQEAVKTFISLVDKLADMR 185

Query: 143 RKALVRQTTPSTGG 156
            K  ++   P   G
Sbjct: 186 MKEKLKGRPPKKDG 199
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2674

A DNA sequence (GASx976) was identified in *S.pyogenes* <SEQ ID 7847> which encodes the amino acid sequence <SEQ ID 7848>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2478(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC79545 GB:U88974 ORF30 [Streptococcus thermophilus temperate
           bacteriophage 01205]
```

```
 Identities = 43/119 (36%), Positives = 66/119 (55%), Gaps = 16/119 (13%)

Query: 9   SKEILHNLDYEAISVTLDSNKIG-----KKVVPAGTILAGKDKSIFEDRKQKVETVINEE 63
           +  I+ +L Y+A+S T+DS+  G      KK + AGT++AG    SIF+DR + V
Sbjct: 9   TSNIVRSLPYKAVSATVDSSYPGVLVDGKKYIKAGTLVAGNGGSIFDDRTKSV------- 61

Query: 64  VSTKEYVDGILLLTDVDLINGDAVGSCVYRGTINADKLADSSVAENYDDLEEVLPHIVPI 122
           V  K   +GI+L DVDLI  + V S +Y G +  DK+    +    D +++ LP + PI
Sbjct: 62  VENKTEPEGIVLYDVDLTIDNTV-SVLYAGEVYKDKVNGGDIT---DTVKKALPLVKPI 116
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2675

A DNA sequence (GASx978) was identified in *S.pyogenes* <SEQ ID 7849> which encodes the amino acid sequence <SEQ ID 7850>. Analysis of this protein sequence reveals the following:

```
Possible site: 60

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4238(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC79546 GB:U88974 ORF31 [Streptococcus thermophilus temperate
            bacteriophage O1205]
 Identities = 195/343 (56%), Positives = 256/343 (73%), Gaps = 1/343 (0%)

Query: 1   MALIHEIITSENIKGFYNAKNENVENTLGEKAPPPKQQLGLKLSPIKGAAGKPVTLKAAA 60
           M LI++ +T+ NI G++NA  ENV +TLGE  FP ++QLG KLS+IKGA+G+ V LKAAA
Sbjct: 1   MSLIYDKVTASNIAGYFNALQENVSSTLGESIFPARKQLGTKLSYIKGASGQSVALKAAA 60

Query: 61  FDTKVPLRDRMAVELIDEEMPFFKEAMLVKEADRQQLNMLAQTKNMELIDTILASIYNDQ 120
           FDT V +RDR++  E+ DE+MPFFKEAMLVKE DRQQLN++  + N  L++TI+A I+ND
Sbjct: 61  FDTNVTIRDRVSAEMHDEQMPFFKEAMLVKENDRQQLNLVKDSGNAVLVNTIVAGIFNDN 120

Query: 121 ATLIAGAKARLEAMRMEVLSKGKIHIQSNGVMKDIDYGLAEDQTTKPDAKWDSAGTATPL 180
            TL+ GA+ARLEAMRM+VL+ GKI   S+GV KDIDYG+  D   +    W  G ATPL
Sbjct: 121 LTLVNGARARLEAMRMQVLATGKIAFTEDGVNKDIDYGVKPDHKKQVSKSWAEPG-ATPL 179

Query: 181 KDIEKAIEKMAERGFVPEAIIMNSKTFSLIKNAESTLDVVKPNAPNGAAVTKRDLNTYLE 240
            D+E AIE  E G  PE  +MN+KTF LI+ A ST+ V+KP+A +G+AVTK +L  Y+
Sbjct: 180 ADLEDAIETARELGLNPERAVMNAKTFGLIRKAASTVKVIKPLAGDGSAVTKAELENYIA 239

Query: 241 DELQIKVILKDGMFVGDDGESRKYFPDGPATLVPNGNIGYTVPGTPESQSDLLGGEATDA 300
            D   + ++L++G +  D GE  K++PDG  TL+PNG LG TVPGTPTPESDLL      +A
Sbjct: 240 DNFGVSIVLENGTYRNDKGKVSKFYPDGHLTLIPNGPLGNTVPGTPTPESDLLFADNTVNA 299

Query: 301 NVSIVETGIAITTTKTTDPVNVQTKVSMIALPSFERLEEVHII 343
            V IV+ GIA+TTTKTTDPVNVQTKVSM+ALPSFERL++V+++
Sbjct: 300 EVEIVDNGIAVTTTKTTDPVNVQTKVSMVALPSFERLDDVYML 342
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2676

A DNA sequence (GASx979) was identified in *S.pyogenes* <SEQ ID 7851> which encodes the amino acid sequence <SEQ ID 7852>. Analysis of this protein sequence reveals the following:

```
Possible site: 46

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3319(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2677

A DNA sequence (GASx980) was identified in *S.pyogenes* <SEQ ID 7853> which encodes the amino acid sequence <SEQ ID 7854>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2385(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC34404 GB:AP158600 gp113 [Streptococcus thermophilus
           bacteriophage Sfi11]
 Identities = 53/109 (48%), Positives = 79/109 (71%), Gaps = 4/109 (3%)

Query: 11  IVKNVKLDLGIEDDNQDQLLEMLLNRITDHFKANYGVLEIDNAPSFVLEDCLIARFNRRG 70
           ++++NV +DL I DDN   LL +LL RI +HFKA YGV E+D+  +F+ EDCL+ RFNRRG
Sbjct: 9   VIQNVSVDLNINDDN---LLGILLERIVNHFKAEYGVDEVDDNLAPIPEDCLVKRFNRRG 65

Query: 71  SKRAKTEKVSGHKTTYYDHLNKFPYDAMIMAKLNLIKDKSRKGGLYPL 119
           +E A++E ++GH  +YYD+ NEF+PYD M+  +L      ++++G + PL
Sbjct: 66  AEGARSESIDGHSMSYYDNENEFDPYDNMLQ-RLYGTSGQAKEGEVLPL 113
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2678

A DNA sequence (GASx981) was identified in *S.pyogenes* <SEQ ID 7855> which encodes the amino acid sequence <SEQ ID 7856>. Analysis of this protein sequence reveals the following:

```
Possible site: 49

>>> Seems to have no N-terminal signal sequence
```

```
----- Final Results -----
               bacterial cytoplasm --- Certainty=0.5714(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA59188 GB:X84706 b3 [Bacteriophage B1]
 Identities = 28/82 (34%), Positives = 49/82 (59%), Gaps = 2/82 (2%)

Query: 1  MRYADRVTFVKTT-DEQYNPDLGEYTHTEVISITKPCFVMDMGMEKSVQIFGDYQKDRKV 59
          +RY D VTF+K + D  Y+PDLGE+   E          + D+G ++SV++FGD +K  KV
Sbjct: 1  IRYLDEVTFIKESPDSHYDPDLGEWVEKEPTRTVFSANITDIGTDRSVEVFGDIKKGAKV 60

Query: 60 IYLKQFYT-KAPDYCEYEGRRY 80
          + +   +    +DY E++ +++
Sbjct: 61 MRMMPLPNMPKYDYIEFDNKKW 82
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2679

A DNA sequence (GASx982) was identified in *S.pyogenes* <SEQ ID 7857> which encodes the amino acid sequence <SEQ ID 7858>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2509(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC34406 GB:AF158600 gp114 [Streptococcus thermophilus
           bacteriophage Sfi11]
 Identities = 44/103 (42%), Positives = 65/103 (62%), Gaps = 5/103 (4%)

Query: 17  GLKKKLELIIKKDAVKK---IVRDNGTQLQRKMINKAVPTKGYSTGAIRRSITMQIGDGG 73
           GL +  + ++K  + +K   ++R  G++L+    +N+A F KGYSTGATRRSIT+Q+
Sbjct: 8   GLDEMAQSLLKNASPEKRSKVLRKYGSKLKEAAVNRAQFNKGYSTGATRRSITLQVESDK 67

Query: 74  LSVKVKPGTHYAGYLERGTRLMSKQDPVLPALKEQKVKFRKDL 116
           +V+    T Y+GYLE GTR M  QPP+ PAL E   K  ++L
Sbjct: 68  ATVEAL--TGYSGYLEVGTRKMEAQPPMKPALDEVAPKMVEEL 108
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2680

A DNA sequence (GASx983) was identified in *S.pyogenes* <SEQ ID 7859> which encodes the amino acid sequence <SEQ ID 7860>. Analysis of this protein sequence reveals the following:

```
Possible site: 45
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3098(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA32612 GB:L31366 putative [Bacteriophage Tuc2009]
 Identities = 88/129 (68%), Positives = 108/129 (83%)

Query: 1   MIKTRDQSIFDEMFKRIQSLGFKVYDYKPMTEVPYPFVEMESTDASYIPNKDDIKGSVSL 60
           MIKTRDQSIFDE+FKRIQ+LG+ VYDYKPM EV YPFVE+E+T   +  NK DIKG+V L
Sbjct: 1   MIKTRDQSIFDELFKRIQALGYTVYDYKPMNEVGYPFVELENTQTIHEANKTDIKGTVSL 60

Query: 61  MLSVWGVQKKRKQVSDMASAIPSQALTVESSDVFRWSLNTRQSSIQMLDDTTTVTPLKRA 120
            LSVWG+QKKRK+VSDMAS IP+QAL + ++D + W+LN++ S+IQMLDDTTT TPLKRA
Sbjct: 61  SLSVWGLQKKRKEVSDMASNIPNQALNISATDGYSWALNSQASTIQMLDDTTTHTPLKRA 120

Query: 121 IVTLRPNLR 129
           ++ L F LR
Sbjct: 121 LINLEFRLR 129
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2681

A DNA sequence (GASx984R) was identified in *S.pyogenes* <SEQ ID 7861> which encodes the amino acid sequence <SEQ ID 7862>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1736(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2682

A DNA sequence (GASx985) was identified in *S.pyogenes* <SEQ ID 7863> which encodes the amino acid sequence <SEQ ID 7864>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA32613 GB:L31366 structural protein [Bacteriophage Tuc2009]
 Identities = 81/185 (43%), Positives = 111/185 (59%), Gaps = 22/185 (11%)

Query: 4   QLRAKQGIHSILLPWLLKEASSEAATKLAPQTEHEVGKSRDVDGQKTKDGIIQSVGALEY 63
           +L AKQG   ILL+RLL +A+ EAA KLAPQTEH    K+RD +    TKDG I S+ A+EY
Sbjct: 3   ELTAKQGKDILLLTKLLSKATKEAANKLAPQTEHSNSKTRDYNTTATKDGTIGSLAAIEY 62

Query: 64  DFKATSILAKGDVLAAKLEKAMENGELVEIWDIDLEETSKNGDSDNKLANVWGIDKNSTN 123
              ATSI A GD    +++KA ++GE++++W+ID  E
Sbjct: 63  SLSATSIAANGDPHLDSMDKAFDDGSIIDVWEIDKAEKG---------------------- 101

Query: 124 RGNGKYLATYYQGYISSFSAKKNAEENIEIEMEFAINGVGQKGFATLTDAQKAAVQYAFK 183
             +GKY A Y + Y++SFS + N+E+ +E+ +EF + G  QKG ATLT+ Q   VQY FK
Sbjct: 102 -SDGKYKAKYLRAYLTSFSYEPNSFDALELSLEFGVFGKPQKGQATLIEEQANVVQYVFK 160

Query: 184 DTTKG 188
             DT  G
Sbjct: 161 DTVAG 165
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2683

A DNA sequence (GASx986) was identified in *S.pyogenes* <SEQ ID 7865> which encodes the amino acid sequence <SEQ ID 7866>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2273(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA59132 GB:X84706 a2 [Bacteriophage B1]
 Identities = 54/111 (48%), Positives = 72/111 (64%), Gaps = 1/111 (0%)

Query: 1   MQLEIKGKTHNVKFGTRFVAEMDKGHIAERQGFKFGAGLQGSV-PFLIDHSVVTLAEVIY 59
           M+L IKGK  + KFG +FV E+DKN + E+ G FG  L   + P L   ++ TL+ V++
Sbjct: 1   MELTIKGKQVHFKFGVKFVRELDKNLVIEQNGVSFGLALAVKIIPELEMANIATLSNVLF 60

Query: 60  TGTITKPPRPRSLNDIYDYIDEVRDIEKLFDDVLDELRQSNASKLPMAQVRK 110
             G  TE P+ S  DI D+IDE EDIEKLFDDVL E+ +SN  KL  A++ K
Sbjct: 61  LGNRTETPKLSQGDIDDFIDECEDIEKLFDDVLKEITESNIGKLIKAMTK 111
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2684

A DNA sequence (GASx987) was identified in *S.pyogenes* <SEQ ID 7867> which encodes the amino acid sequence <SEQ ID 7868>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2735(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA59193 GB:X84706 c2 [Bacteriophage B1]
 Identities = 40/111 (36%), Positives = 57/111 (51%), Gaps = 10/111 (9%)

Query: 2   IVLNCIRYLGMTDINEIGRLTLYEYDLLMTGKALAAVDESHKAHKQAWINHQVTATKLVG 61
           +++  +R  G+ D++   R+T+ EY +      L  +DE    ++QAW N QV ATK  G
Sbjct: 15  MKIRFLRCPGIQDLSVFERMTIREYSIRSIAFQLRTLDEEEFIYEQAWANWQVQATKQQG 74

Query: 62  GKDNKKEVPVVKKFKDFD---YEEEIRKI-TQEIDRGYDKKGMDLLLKAN 108
            K      P+Y   FK FFD     E KI  I + E        D K +DL+ KAN
Sbjct: 75  KK------PLYPFKKFFDKKKLENSILGIESPENKFKDNNKLIDLMKKAN 119
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2685

A DNA sequence (GASx989) was identified in *S.pyogenes* <SEQ ID 7869> which encodes the amino acid sequence <SEQ ID 7870>. Analysis of this protein sequence reveals the following:

```
Possible site: 60

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2869(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA66560 GB:X97918 gene 19.1 [Bacteriophage SPP1]
 Identities = 66/232 (28%), Positives = 106/232 (45%), Gaps = 12/232 (5%)

Query: 38  FRTLTVSGRDVVDLEHQTTSVLGRNGEYPENATVEVRKLEIKAKISGKDNKS-MRLQYEK 96
           F   V GR V  +E    ++ G +G      ++  R+LE+ A + G   ++ +R + E
Sbjct: 24  FLVQEVRGRSVYSIEMGKRTIAGVDGGVITTRSLPARELEVDAIVPGDGTETDLRRRIEY 83

Query: 97  INKLIVSHNQVPLSFSDEPDRNYLGIFKSKDVPKEVSNEQIIGLTPICYNPPK-----MS 151
           LN L+      V ++FSDEP R Y G ++      +E      + L P C +P K      +
Sbjct: 84  LNFLLHRDTDVPITFSDEPSRTYYGRYEFATEGDEKGGFHKVTLNFYCQDPLKYGPEVTT 143

Query: 152 DVKTKKGTSIQNSGLFQTKPIITLNLSSPTKEIKLLHVESQKYIRLT----GTYTTDEIK 207
           DV T   T ++N GL  T P I    S+   E ++  ++     ++      G  T D +
Sbjct: 144 DV-TTASTPVKNTGLAVTNPTIRCVFSTSATEYEMQLLDGSTVVKFLKVKYGFNTGDTLV 202

Query: 208 IDMATGKITQNGRNILGDLDMINSRYFELLPGNNTLQCANAATAEFREVVL 259
           ID     +T NG++I+  L +I S + +L P  NT   A    T  F E +L
Sbjct: 203 IDCHERSVTLNGQDIMPALTLIQSDWIQLKPQVNTYLKATQPSTIVFTEKFL 253
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2686

A DNA sequence (GASx990) was identified in *S.pyogenes* <SEQ ID 7871> which encodes the amino acid sequence <SEQ ID 7872>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2861(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:EAB04681 GB:AP001510 unknown conserved protein in others
            [Bacillus halodurans]
 Identities = 116/449 (25%), Positives = 198/449 (43%), Gaps = 79/449 (17%)

Query: 2   IYLFDKLERLVATVG-TDDLLSWHFKVKNNDWDQASPEVPVDYDVEPPVYPGPFNYDPHQ 60
           +++FD+ ++L+ T+  +  L+   F+ + N      F   ++   E   +       + HQ
Sbjct: 4   LFIFDREDQLLTILTESTGLVRALFREELNRVPNQPFAFTIEASSEEAKHV----IEEHQ 59

Query: 61  -----KEDVFKLFKVDIYNLEDSKFYKG------LDKASDLDTIAIIKDKRFDQSGADA 109
                KE    +LF + +   LED     G      + A  +L    I++      Q + +A
Sbjct: 60  VVFRDKEGDLRLFVIKE--LEDVDGLDGPQTTAICEPAPMELAKHMIVEQSVVNQPAHEA 117

Query: 110 CIDGALEGTGYQVGKVEGITNVRTLSYYYISPRAALIKIVEAPNCEFNVRYTF-INNKIT 168
            ++ AL+GT +  G VE       T  + Y+S   A+  I+  +F    TF   N+IT
Sbjct: 118 -LNVALQGTRW-TGSVEVNLGNATEHFSYVSAIEAVWNILVTWGGDFKDVVTFNAENRIT 175

Query: 169 SRYIDLKKRFGKPTGKQFEHGNNLLKVVYEESTDDIVTCLIGRGKGEEIQHEEAEPKDVE 228
           S  I + +R G   GK+FE  +N+ + +        VT L GRG    +Q E  E    +
Sbjct: 176 SHQIKIVQRRGVDRGKRFEIDHNI-EQIERTILSYPVTALYGKGAS--LQGENGE----D 228

Query: 229 GHLPQEERRQGVGRRIEFTDVVWSVEKGDPIDKPAGQNFVALDGAREEYGLQNGELKHR 286
            G L          +F +V W   G P+DKP GQ +V    A ++YG   NG+L HR
Sbjct: 229 GSL-------------DFGEVEWRKSAGAPVDKPKGQLMVGDPEALQKYGRKHNGQLLHR 275

Query: 289 WGVPVNEEIEDKTELLKATWEELQRLSIPIRIYKAEILDIGPETWKGDSVAIIYDEVKIA 348
            G+F N  IED  ELL+ TWE+LQ+ S P   Y+  +                +++ +
Sbjct: 276 EGIPQNTNIEDPEKLLKKTWKQLQKSSKPEVHYRLSVR--------------LFEHIS-- 319

Query: 349 FETRVDEIDIDKLNFNRSVVTLGDYSVVQNR------RSRSRKEAVQ-NMIDESLETTTD 401
                      +     +LGD ++  +R       E +SR  A++ +++D     + +
Sbjct: 320 -------------GYEHEQASLGDTAIAIDRQFSRPIEIQSRIIAIRYDLVDIDQTGMVE 366

Query: 402 LGMTFQEFLQSIKKRIETGKKEMEDNWRK 430
            +G    L G++++R+E   +E+E N  K
Sbjct: 367 MCQFLS--LNGMDERLERIIESIEKNQGK 393
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2687

A DNA sequence (GASx991) was identified in *S.pyogenes* <SEQ ID 7873> which encodes the amino acid sequence <SEQ ID 7874>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2584(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA98101 GB:M19348 hyaluronidase [Streptococcus pyogenes phage
            H4489A]
 Identities = 314/371 (84%), Positives = 338/371 (90%), Gaps = 1/371 (0%)

Query: 1   MAENIPLRVQPKRMKAAKWASSDVVLLHGEIGFETDTGPAKPGDGQNTFSKLKYLTGPKG 60
            M ENIPLRVQPKRM A EWA SDV+LLHGEIGFETDTGPAKPGDGQNTFSKLKYLTGPKG
Sbjct: 1   MTENIPLRVQPKRMSADEWARSDVILLHGEIGFETDTGPAKPGDGQNTFSKLKYLTGPKG 60

Query: 61  PKGDTGLQGKTGGTGSRGPAGKPGTTDVDQLQNKPDLGAFAQKEETNSKITKLESSKADK 120
            PKGDTGLQGKTGGTG RGPAGKPGTTDVDQLQNKPDLGAFAQKEETNSKITKLESSKADK
Sbjct: 61  PKGDTGLQGKTGGTGPRGPAGKPGTTDVDQLQNKPDLGAFAQKEETNSKITKLESSKADK 120

Query: 121 NAVVLKAESNAKLDEKLNLKGGVMTGQLQFKPN-SGIKPSSSVGGAINIDMSKSEGAAMV 179
            +AVV KAES  +LD+KL+L GG++TGQLQFKPN SGIKPSSSVGGAINIDMSKSEGAAMV
Sbjct: 121 SAVVSKAESKIKLDKKISLTGGIVTGQLQFKPNKSGIKPSSSVGGAINIDMSKSEGAAMV 180

Query: 180 MYTNKDTTDGPLMILRSNKDTFDQSVQFVDYKGTTNAVNIVMRQPTTPNFSSALNITSAN 239
            MYTNKDTTDGPLMILRS+KDTFDQS QFVDY G TNAVNIVMRQP+ PNFSSALNITSAN
Sbjct: 181 MYTNKDTTDGPLMILRSDKDTFDQSAQFVDYSGKTNAVNIVMRQPSANFSSALNITSAN 240

Query: 240 EGGSAMQIRGVEKALGTLKITHENPSVDKEYDKNAAALSIDIVKKQKGSKGTAAQGIYIN 299
            EGGSAMQIRGVEKALGTLKITHENP+V+ +YD+NAAALSIDIVKKQKGSKGTAAQGIYIN
Sbjct: 241 EGGSAMQIRGVEKALGTLKITHENPNVEAKYDENAAALSIDIVKKQKGSKGTAAQGIYIN 300

Query: 300 STSGTTGKLLRIRNLNDDKFYVKPDGGFYAKETSQIDGNLKLKDPIANDHAATKAYVDGE 359
            STSGT GK+LRIRN N+DKFYV PDGGF++    S + GNL +KDP +  HAATK YVD +
Sbjct: 301 STSGTAGKMLRIRNKNEDKFYVGPDGGFHSGANSTVAGNLTVKDPTSGKHAATKDYVDEK 360

Query: 360 VEKLKALLAAK 370
            + +LK L+  K
Sbjct: 361 IAELKKLILKK 371
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2688

A DNA sequence (GASx993) was identified in *S.pyogenes* <SEQ ID 7875> which encodes the amino acid sequence <SEQ ID 7876>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1358(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2689

A DNA sequence (GASx995) was identified in *S.pyogenes* <SEQ ID 7877> which encodes the amino acid sequence <SEQ ID 7878>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0855(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC34418 GB:AF158600 gp149 [Streptococcus thermophilus
           bacteriophage Sfi11]
 Identities = 27/95 (28%), Positives = 50/95 (52%), Gaps = 2/95 (2%)

Query: 9   KYPQLDGTGAVASTHIIIAAEDGAVIPQLIKQDLTSTNDTEIIKAALEEFKKSEYVEIAM 68
           K  + D +GA  +T +I+   DGA +P  +  +    ++TE++K ALE   +  + + A
Sbjct: 26  KSKEYDASGAAYATKVILKNRDGAYVPVFLPVEKIDLSNTELLKEALEVIYQENPFQRAE 85

Query: 69  GEAVQKVDDLEKISQETAKTAKTAQTAAGLAKVSA 103
            E   ++D  EKI +  A + K  +T A + + S+
Sbjct: 86  NEKFNELD--EKIKEYEALSKKATETIAKMEEASS 118
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2690

A DNA sequence (GASx996) was identified in *S.pyogenes* <SEQ ID 7879> which encodes the amino acid sequence <SEQ ID 7880>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -4.62   Transmembrane    9 -  25 (   7 -  26)

----- Final Results -----
             bacterial membrane --- Certainty=0.2848(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2691

A DNA sequence (GASx997) was identified in *S.pyogenes* <SEQ ID 7881> which encodes the amino acid sequence <SEQ ID 7882>. Analysis of this protein sequence reveals the following:

```
Possible site: 41

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -3.66   Transmembrane   38 -  54 (  35 -  55)

----- Final Results -----
                bacterial membrane --- Certainty=0.2466(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2692

A DNA sequence (GASx998R) was identified in *S.pyogenes* <SEQ ID 7883> which encodes the amino acid sequence <SEQ ID 7884>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -9.87   Transmembrane   47 -  63 (  41 -  72)

----- Final Results -----
                bacterial membrane --- Certainty=0.4949(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2693

A DNA sequence (GASx999) was identified in *S.pyogenes* <SEQ ID 7885> which encodes the amino acid sequence <SEQ ID 7886>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                 bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2694

A DNA sequence (GASx1001) was identified in *S.pyogenes* <SEQ ID 7887> which encodes the amino acid sequence <SEQ ID 7888>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-10.51   Transmembrane   18 -  34 (  16 -  34)

----- Final Results -----
                bacterial membrane --- Certainty=0.5203(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2695

A DNA sequence (GASx1002) was identified in *S.pyogenes* <SEQ ID 7889> which encodes the amino acid sequence <SEQ ID 7890>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -3.61   Transmembrane   12 -  28 (  11 -  33)

----- Final Results -----
                bacterial membrane --- Certainty=0.2444(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein is similar to AF186180 from *S.equi*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2696

A DNA sequence (GASx1003) was identified in *S.pyogenes* <SEQ ID 7891> which encodes the amino acid sequence <SEQ ID 7892>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                 bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein is similar to SeeH from *S.equi*:

```
>GP:AAF72809 GB:AF186180 SeeH [Streptococcus equi]        Length = 236
 Identities = 233/236 (98%), Positives = 234/236 (98%)

Query: 1   MRYNCRYSHIDKKIYSMIICLSFLLYSNVVQANSYNTTNRHNLESLYKHDSNLIEADSIK 60
           MRYNCRYSHIDKKIYSMIICLSFLLYSNVVQANSYNTTNRHNLESLYKHDSNLIEADSIK
Sbjct: 1   MRYNCRYSHIDKKIYSMIICLSFLLYSNVVQANSYNTTNRHNLESLYKHDSNLIEADSIK 60

Query: 61  NSPDIVTSHMLKYSVKDKNLSVFFEKDWISQEFKDKEVDIYALSAQEVCECPGKRYEAPG 120
           NSPDIVTSHMLKYSVKDKNLSVFFEKDWISQEFKDKEVDIYALSAQE CECPGKRYEAPG
Sbjct: 61  NSPDIVTSHMLKYSVKDKNLSVFFEKDWISQEFKDKEVDIYALSAQEACECPGKRYEAPG 120

Query: 121 GITLTNSEKKEIKVPNVWDKSKQQPPMFITVNKPKVIAQEVDIKVRKLLIKKYDIYNNR 180
           GITLTNSEKKEIKVP+NVWDKSKQ PPMFITVNKPKVIAQEVDIKVRKLLIKKYDIYNNR
Sbjct: 121 GITLTNSEKKEIKVPINVWDKSKQHPPMFITVNKPKVIAQEVDIKVRKLLIKKYDIYNNR 180

Query: 181 EQKYSKGTVTLDLNSGKDIVFDLYYPGNGDPNSMLKIYSNNERIDSTQFHVDVSIS 236
           EQKYSKGTVTLDLNSGKDIVFDLYYPGNGDPNSMLKIYSNNERIDSTQFHVDVSIS
Sbjct: 181 EQKYSKGTVTLDLNSGKDIVFDLYYPGNGDPNSMLKIYSNNERIDSTQFHVDVSIS 236
```

There is also homology to a *S.aureus* enterotoxin:

```
>GP:AAA19777 GB:U11702 enterotoxin H [Staphylococcus aureus]
 Identities = 70/215 (32%), Positives = 108/215 (49%), Gaps = 19/215 (8%)

Query: 27  SNVVQANSYNTTNRHNLESLYKHDSNLIEADSI-KNSPDIVTSHMLKYSVKDKNLSVFFE 85
           +++  AN+Y   N   ++   K D    E D I +N  D      +K++  D
Sbjct: 34  TDLALANAYGQYNHPFIKENIKSDEISGEKDLIPRNQGDSGNDLRVKFATAD-------- 85

Query: 86  KDWISQEFKDKEVDIYALSAQEVCECPGKRYEA--PGGITLTNSEK--KEIKVPNVWDK 141
              ++Q+FK+K VDIY  S    CE    +    +GG TL NSEK  +E  +  NVW
Sbjct: 86  ---LAQKFGNKNVDIYGASFYYKCEKISENISECLYGGTTL-NSEKLAQERVIGANVWVD 141

Query: 142 SKQQPPMFITVNKPKVTAQEVDIKVRKLLIKKYDIYNNREQKYSKGTVTLDLNSGKDIVF 201
             Q+    I  NK  VT QE+DIK+RK+L  KY IY  ++ + SKG +  D+ ++D  F
Sbjct: 142 GIQKETELIRTNKKNVTLQELDIKIRKILSDKYKIY-YKDSEISKGLIEFDMKTPRDYSF 200

Query: 202 DLYYPGNGDPNSMLKIYSNNERIDSTQF-HVDVSI 235
           D+Y     +   + KIY +N+ + S    H+DV++
Sbjct: 201 DIYDLKGRNDYEIDKIYENKTLKSDDISHIDVNL 235

>GP:AAC26661 GB:AF064774 extracellular enterotoxin type I precursor
          [Staphylococcus aureus]
 Identities = 68/214 (31%), Positives = 109/214 (50%), Gaps = 27/214 (12%)

Query: 42  NLESLY-KHDSNLIEADSIKNSPDIVTSHMLKYSVKDKNLSVFFEKDWIS-QEFKDKEVD 99
           NL + Y KHD   ++  + KN P    ++ L++S    +L +    +W    +FK K++D
Sbjct: 32  NLRNFYTKHDYIDLKGVTDKNLP---IANQLEFSTGTNDL-ISRSNNWDEISKFKGKKLD 87

Query: 100 IYALSAQEVCECPGKRYEAPGGITLTNSEKKEI-KVPVNVWDKSKQQPPMF--ITVNKPK 156
           I+ +    C    K    +GG TL+         K+P+N+W   K +       I  NK
Sbjct: 88  IFGIDYNGPC----KSKYMYGGATLSGQYLNSARKIPINLWVNGKHKTISTDKIATNKKL 143

Query: 157 VTAQEVDIKVRKLLIKKYDIYNNRE----------QKYSKGTVTLDLNSGKDIVFD 202
           VTAQE+D+K+R+ L ++Y+IY +                ++ G V   LN+ K   +D
Sbjct: 144 VTAQSIDVKLRRYLQEKYNIYGHNNTGKGKEYGYKSKFYSGFNNGKVLFHLNNEKSFSYD 203

Query: 203 LYYPGNGDPNSMLKIYSNNERIDSTQFHVDVSIS 236
           L+Y G+G   S LKIY +N+ I+S +FH+DV IS
Sbjct: 204 LFYTGDGLPVSFLKIYEDNKIIESEKFHLDVEIS 237

>GP:AAC28968 GB:U93688 enterotoxin [Staphylococcus aureus]
 Identities = 70/244 (28%), Positives = 127/244 (51%), Gaps = 27/244 (11%)

Query: 12  KKIYSMIICLSFLLYSNVVQANSYNTTNRHNLESLYKHDSNLIEADSIKNSPDIVTSHML 71
```

```
           KK+ S+++ ++ ++       A++        NL + Y    + ++    +K++ D    ++ L
Sbjct: 2   KKLISILL-INIIILGVSNNASAQGDIGIDNLRNFYTK-KDFVDLKDVKDN-DTPIANQL 58

Query: 72  KVSVKDKNLSVFPEKDWIS-QEFKDKEVDIYALSAQEVCEKGKRYEAFGGITLTNSE-K 129
           ++S +   +L +   KD+      FK K++D++ +S    C       +Y  +GG+T TN
Sbjct: 59  QFSNESYDL-ISSSKDFNKFSNFKGKKLDVPGISYNGQCNT---KY-IYGGVTATNEYLD 113

Query: 130 KEIKVFVNVW--DKSKQQPPMPITVNKPKVTAQEVDIKVRKLLIKKYDIYNNREQK---- 183
           K   +P+N+W      K       ++ NK  VTAQE+D+K+RK L ++Y+IY +   K
Sbjct: 114 KSRNIPINIWINGNHKTISTNKVSTNKKLVTAQEIDVKLRKYLQEEYNIYGHNGTKKGEE 173

Query: 184 ----------YSKGTVTLDLNSGKDIVFDLYYFG-NGDFNSMLKIYSNNERIDSTQFHVD 232
                     ++ G VI   LN+      +DL+Y G +G    S LKIY +N+ ++S +FH+D
Sbjct: 174 YGHKSKFYSGFNIGKVTFHLNNNDTPSYDLFYTGDDGLFKSFLKIYEDNKTVRSEKPHLD 233

Query: 233 VSIS 236
           V IS
Sbjct: 234 VDIS 237
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2697

A DNA sequence (GASx1004R) was identified in *S.pyogenes* <SEQ ID 7893> which encodes the amino acid sequence <SEQ ID 7894>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -2.18   Transmembrane   12 -  28 (  12 -  28)

----- Final Results -----
                bacterial membrane --- Certainty=0.1871(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2698

A DNA sequence (GASx1009) was identified in *S.pyogenes* <SEQ ID 7895> which encodes the amino acid sequence <SEQ ID 7896>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.6391(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2699

A DNA sequence (GASx1011) was identified in *S.pyogenes* <SEQ ID 7897> which encodes the amino acid sequence <SEQ ID 7898>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4528(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2700

A DNA sequence (GASx1024) was identified in *S.pyogenes* <SEQ ID 7899> which encodes the amino acid sequence <SEQ ID 7900>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2701

A DNA sequence (GASx1033) was identified in *S.pyogenes* <SEQ ID 7901> which encodes the amino acid sequence <SEQ ID 7902>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1652(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2702

A DNA sequence (GASx1039) was identified in *S.pyogenes* <SEQ ID 7903> which encodes the amino acid sequence <SEQ ID 7904>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.06   Transmembrane   15 -  31 (  15 -  31)

----- Final Results -----
                bacterial membrane --- Certainty=0.1426(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2703

A DNA sequence (GASx1058) was identified in *S.pyogenes* <SEQ ID 7905> which encodes the amino acid sequence <SEQ ID 7906>. Analysis of this protein sequence reveals the following:

```
Possible site: 60

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.5484(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2704

A DNA sequence (GASx1077) was identified in *S.pyogenes* <SEQ ID 7907> which encodes the amino acid sequence <SEQ ID 7908>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.4848(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2705

A DNA sequence (GASx1080) was identified in *S.pyogenes* <SEQ ID 7909> which encodes the amino acid sequence <SEQ ID 7910>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-12.42   Transmembrane  107 - 123 (  93 - 133)
   INTEGRAL    Likelihood =-11.20   Transmembrane   20 -  36 (  14 -  44)
   INTEGRAL    Likelihood = -8.39   Transmembrane  226 - 242 ( 218 - 246)
   INTEGRAL    Likelihood = -5.52   Transmembrane  129 - 145 ( 126 - 148)
   INTEGRAL    Likelihood = -4.46   Transmembrane  160 - 176 ( 159 - 183)
   INTEGRAL    Likelihood = -1.44   Transmembrane   55 -  71 (  55 -  72)

----- Final Results -----
               bacterial membrane --- Certainty=0.5967(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2706

A DNA sequence (GASx1081) was identified in *S.pyogenes* <SEQ ID 7911> which encodes the amino acid sequence <SEQ ID 7912>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-13.00   Transmembrane  103 - 119 (  91 - 129)
   INTEGRAL    Likelihood =-11.46   Transmembrane  208 - 224 ( 203 - 230)
   INTEGRAL    Likelihood = -8.28   Transmembrane   54 -  70 (  46 -  71)
   INTEGRAL    Likelihood = -5.79   Transmembrane  160 - 176 ( 155 - 181)
   INTEGRAL    Likelihood = -4.25   Transmembrane  127 - 143 ( 125 - 149)

----- Final Results -----
               bacterial membrane --- Certainty=0.6201(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2707

A DNA sequence (GASx1089) was identified in *S.pyogenes* <SEQ ID 7913> which encodes the amino acid sequence <SEQ ID 7914>. Analysis of this protein sequence reveals the following:

```
Possible site: 37
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2999(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2708

A DNA sequence (GASx1109) was identified in *S.pyogenes* <SEQ ID 7915> which encodes the amino acid sequence <SEQ ID 7916>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1270(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2709

A DNA sequence (GASx1114R) was identified in *S.pyogenes* <SEQ ID 7917> which encodes the amino acid sequence <SEQ ID 7918>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4021(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2710

A DNA sequence (GASx1149) was identified in *S.pyogenes* <SEQ ID 7919> which encodes the amino acid sequence <SEQ ID 7920>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -1.70   Transmembrane   12 -  28 (  12 -  29)

----- Final Results -----
                bacterial membrane --- Certainty=0.1680(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2711

A DNA sequence (GASx1150) was identified in *S.pyogenes* <SEQ ID 7921> which encodes the amino acid sequence <SEQ ID 7922>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                 bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2712

A DNA sequence (GASx1160) was identified in *S.pyogenes* <SEQ ID 7923> which encodes the amino acid sequence <SEQ ID 7924>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -3.19   Transmembrane   15 -  31 (  15 -  31)

----- Final Results -----
                bacterial membrane --- Certainty=0.2275(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2713

A DNA sequence (GASx1167) was identified in *S.pyogenes* <SEQ ID 7925> which encodes the amino acid sequence <SEQ ID 7926>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1404(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB99233 GB:U67563 oxaloacetate decarboxylase alpha chain (oadA)
             [Methanococcus jannaschii]
 Identities = 250/453 (55%), Positives = 325/453 (71%), Gaps = 7/453 (1%)

Query: 13  VAITETVLRDGHQSLMATRLSIEDMLPVLTILDKIGYYSLECWGGATFDACIRFLNEDPW 72
            V I +T  RD  QSL+ATR+  EDMLP+    +D++G+YS+E WGGATFDACIR+LNEDPW
Sbjct: 2   VKIVDTTFRDAQQSLIATRMRTEDMLPIAEKMDEVGFYSMEVWGGATFDACIRYLNEDPW 61

Query: 73  ERLRTLKKGLPWTRLQMLLRGQNLLGYRHYADDIVDKFISLSAQNGIDVFRIFDALNDPR 132
            ERLR LKK + NT LQMLLRGQNL+GYRHY DDIV+KF+  + +NGID+FRIFDALND R
Sbjct: 62  ERLRALKKRINTPLQMLLRGQNLVGYRHYPDDIVEKFVIKAHENGIDIFRIFDALNDVR 121

Query: 133 NIQQALRAVKKTGKEAQLCIAYTTSPVHTLNYYLSLVKELVEMGADSICIKDMAGILTPK 192
            N++ A++  KK G E Q  I YT SPVHT++ Y+ L K+L EMG DSICIKDMAG+LTP
Sbjct: 122 NMETAIKTAKKVGAEVQGAICYTISPVHTIDQYVELAKKLEEMGCDSICIKDMAGLLTPY 181

Query: 193 AAKELVSGIKAMTNLPLIVHTHATSGISQMTYLAAVEAGADRIDTALSPFSEGTSQPATE 252
                ELV  +K   +LP+ VH+H TSG++ MTYL  +AGAD +D A+SPF+ GTSQP TE
Sbjct: 182 EGYELVKRLKEEISLPIDVHSHCTSGLAPMTYLKVIEAGADMVDCAISPFAMGTSQPPTE 241

Query: 253 SMYLALKEASYDITLDETLLEQAANHLRQARQKYLADGILDPSLLFPDPRTLQVQVPGGM 312
            S+ +ALK   YD  LD  LL +  ++  + R+KY    +  P     D R L YQVPGGM
Sbjct: 242 SIVVALKGTKYDTGLDLKLLNEIRDYFMKVRKKYKM--LFSPISQIVDARVLVYQVPGGM 299

Query: 313 LSNMLSQLKQANAESKLEEVLAEVPRVRKDLGYPPLVTPLSQMVGTQAAMNVLLGKPYQM 372
            LSN++SQLK+  A  K EEVL E+PRVRKDLGYPPLVTP SQ+VGTQA +NV+  + Y++
Sbjct: 300 LSNLVSQLKEQGALDKFEEVLQEIPRVRKDLGYPPLVTPTSQIVGTQAVLNVLTERYKI 359

Query: 373 VSKEIKQYLAGDYGKTPAPVNEDLKRSQI--GSADVTINRPADQLSPEFEVLK--AEVAD 428
            ++ E+  Y+ G YGK PAP+N +L +  +  G  P+T  RPAD L PE+E +K  AE
Sbjct: 360 ITNEVVNYVKGFYGKPPAPINPELLKRVLDEGKKPITC-RPADLLPPEWEKVKKEAEEKG 418

Query: 429 LAQTDRDVLTYALFPSVAKPFLTTKYQTDDVIK 461
            + + +ED+LTYALFP +A  FL  + + + + K
Sbjct: 419 IVKKEEDILTYALYPQIAVKFLRGELKAEDIPK 451
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2714

A DNA sequence (GASx1168) was identified in *S.pyogenes* <SEQ ID 7927> which encodes the amino acid sequence <SEQ ID 7928>. Analysis of this protein sequence reveals the following:

```
Possible site: 38

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL     Likelihood = -7.11   Transmembrane   16 -  32 (   2 -  34)

----- Final Results -----
                bacterial membrane --- Certainty=0.3845(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2715

A DNA sequence (GASx1170) was identified in *S.pyogenes* <SEQ ID 7929> which encodes the amino acid sequence <SEQ ID 7930>. Analysis of this protein sequence reveals the following:

```
Possible site: 51

>>> Seems to have no N-terminal signal sequence
   INTEGRAL     Likelihood = -7.06   Transmembrane  211 - 227 ( 208 - 238)
   INTEGRAL     Likelihood = -5.84   Transmembrane  117 - 133 ( 110 - 136)
   INTEGRAL     Likelihood = -5.36   Transmembrane  256 - 272 ( 253 - 274)
   INTEGRAL     Likelihood = -4.67   Transmembrane   44 -  60 (  41 -  64)
   INTEGRAL     Likelihood = -4.19   Transmembrane  287 - 303 ( 287 - 306)
   INTEGRAL     Likelihood = -3.77   Transmembrane  358 - 374 ( 357 - 375)
   INTEGRAL     Likelihood = -2.18   Transmembrane   20 -  36 (  16 -  38)
   INTEGRAL     Likelihood = -0.85   Transmembrane   90 - 106 (  90 - 106)
   INTEGRAL     Likelihood = -0.53   Transmembrane  165 - 181 ( 164 - 181)

----- Final Results -----
                bacterial membrane --- Certainty=0.3824(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA05140 GB:AJ002015 methylmalonyl-CoA decarboxylase,
          beta-subunit [Propionigenium modestum]
 Identities = 231/395 (58%), Positives = 293/395 (73%), Gaps = 19/395 (4%)

Query: 1   MLDVLNQMVQSSGLAHLTVNNLIMICLASFFLVLGIKKEYEPFLMVPIAFGILLVNLPMA 60
           ML  +      S+G   L + ++IM+ +A  FLYL I KE+EP L+VPI+FGILL NLP A
Sbjct: 1   MLQAILDPYHSTGFGLNMGSIIMMLVACVFLYLAIAKESPEPLLLVPISFGILLTNLPFA 60

Query: 61  GLMDHP---------ANG---------NDGGLLYYLYKGTSIGIYPPLIFLCLGASTDPG 102
           G+M  P          A+G           PGGLLYYLY+G  LGI+PPLIFL +GA TDPG
Sbjct: 61  GMMAEPLLEVHEKLSASGAHLYTAHTAENGGLLYYLFQGDHLGIPPPLIFLGVGAMTDPG 120

Query: 103 PLIANPKTILLGNAQVGIFLAPFLAIMLGM-TPQEAASVGIIGGADGPTAIYVTTKLAP 161
           PLI+NPK++LLG AAQ GIF+ FF AI  G+ T QEAAS+GIIGGADGPTAI++++KLAP
Sbjct: 121 PLISNPKSLLLLGAAAQFGIFVTFFGAIASGLFTAQEAASIGIIGGADGPTAIFLSSKIAP 180

Query: 162 DLLSTIALAAAYSYMALVPIIQPPIIKLLTTKAERQVKMTQARTVSQKEKIIFPIMVTIFV 221
            L+  IA++AAYSYMALVPIIQPPI+  LT++ E2++KM+Q R VS++EKIIFPI+VTI V
Sbjct: 181 HLMGPIAVAAYSYMALVPIIQPPIMTALTSEYERKIKMSQLRLVSKREKIIFPIVVTILV 240

Query: 222 SLLVPSATTLVGCLMLGNLVREIKIVPKIVENLQQVVMFCITIILGLTVGAKANGDLFLS 281
```

```
              SL+VP A TLVG LMLGNL RE  +V ++ +  +  ++  ITI LG+TVGA A  + FL
Sbjct: 241 SLIVPPAATLVGMLMLGNLFRECGVVGRLRDTAKNALNIITIFLGVTVGATATAEAFLK 300

Query: 282 ATTLKIIALGLIAFAAGTAGGVLMGKVMVYVLSGNKVMPMIGAAGVSAVPMAARVVQKIGQ 341
            TL I+ LG++AF  GT  GVL+ K M  LS   +NP++G+AGVSAVPMAARV Q +GQ
Sbjct: 301 VETLAILGLGIVAPGIGTGSGVLLAKPMNKLSKEPINPLLGSAGVSAVPMAARVSQVVGQ 360

Query: 342 EEDPSNFLLMHAMGPNVAGVIGSAIASGALLAFPG 376
            + DP+NFLLMHAMGPNVAGVIGSA++ +G LL+ FG
Sbjct: 361 KADPTNFLLMHAMGPNVAGVIGSAVSAGVLLSLFG 395
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2716

A DNA sequence (GASx1171R) was identified in *S.pyogenes* <SEQ ID 7931> which encodes the amino acid sequence <SEQ ID 7932>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.0851(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF93965 GB:AE004165 citG protein [Vibrio cholerae]
 Identities = 100/287 (34%), Positives = 154/287 (52%), Gaps = 12/287 (4%)

Query: 9   ISQLALKALLYEVSLSPKPGLVDRFDNGAHDDMSPITFIDSMIALSPFFQAYIETGFAYA 68
           +  LA  A++ EV L+PKPGLVD  +NGAH DM   TFI S  A++P+  +++  G+  A
Sbjct: 32  VGHLAYHAMMLEVHLTPKPGLVDTANNGAHRDMDLNTFIASAEAIAPYLHSFVSAGWESA 91

Query: 69  KEEPLLLFNRLRQLGQKAEETMFCATQGINTHKGLNFSMALLLGATGAYLARTPHLMTDL 128
                L + LR +G +AE+ MF ATQG+NTHKG+ F + L+ G+ G    A
Sbjct: 92  GNPAAQLLSALRPIGIEAEQAMFAATQGVNTHKGMIFILGLICGSVGWLKANQ------- 144

Query: 129 GRFSKEDTLAICRLVKPMTAHLICQTDLGHLNTKKEPTYGEQLFVTYGIKGPRGEASEGPT 188
              K D   I   ++      L+  +L      +  T GE++  YG+ G RGEA+ G
Sbjct: 145 ---LKIDAQHIGETIRQACQFLVIDELKAKRDCEPETAGERIYRQYGLYGARGEAAEGLA 201

Query: 189 TLTDDALPYFRQMISQN-DPETSQLRLLVYLMSIVEDGNLIHRGGIEAWKGVKAD-MRLL 246
           +  HALP ++  +++     E +     L+ LM+    D NL+ RGG+      V+    +LL
Sbjct: 202 MVMQHALPAYQACLTKGASTEQALWHTLLVLMANNNDSNLVSRGGLAGLHFVQEQAQQLL 261

Query: 247 LQQDLSTTDLRLALSSYNQCLINQHLSPGGSADLLALTFYFAFLEKL 293
            +      ++  AL++ +  LI +HLSPGG+ADLLA T+    L +L
Sbjct: 262 AKGGFLYQEIEQALTALDSVLIEKHLSPGGSADLLAATWLIYELVQL 308
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2717

A DNA sequence (GASx1172R) was identified in *S.pyogenes* <SEQ ID 7933> which encodes the amino acid sequence <SEQ ID 7934>. Analysis of this protein sequence reveals the following:

```
Possible site: 23
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
                bacterial cytoplasm --- Certainty=0.2501(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB12389 GB:Z99107 similar to transcriptional regulator (GntR
             family) [Bacillus subtilis]
 Identities = 60/205 (29%), Positives = 99/205 (48%), Gaps = 3/205 (1%)

Query: 19  PLKIAFYNALKKTIILRQIPVGSRINEKEPSIALNISRTPIRYALGLLSEEHLVEHIPKK 78
           P  + FYN LKK I       G RINE + + +  +SR+PIR A+ LL ++ L++    +
Sbjct: 11  PYYLQFYNQLKKMIFKGTFKPGERINETQLAKSFGVSRSPIREAMRLLEKDGLLKADDRN 70

Query: 79  GIIVKGVSIKDACEIFEIRKALETLATVQAMHLMTEEDFKVMHNLLEDCETFI--AEDDT 136
           G  +  ++ KD  EI++IR  LE LA    +    EE+  ++    LE+ E  I    +DT
Sbjct: 71  GFSITSLTAKDVDEIYKIRIPLEQLAVELVIDEADEEELTTLEKQLEETEKAIHNGTEDT 130

Query: 137 NRILDNFNAFNNLIYSYSQMVRLKEIVIELQAYLVYFRKISISSVERRKRALSKHWMIYR 196
             I   N   F+ L+  +S    LK ++  +   +  + R ++ +   R +  L EH  I+
Sbjct: 131 EIIRLN-QKFHELLVDFSHNRHLKNLLEHVNDLIHFCRILNYTGDHRAETILKEHRRIFE 189

Query: 197 GMKNKDHEQITLITHEHLNSSLEFI 221
             +K K+ E        H N   E +
Sbjct: 190 EVKKKNKEAAKQHVLAHFNHDCEHL 214
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2718

A DNA sequence (GASx1173R) was identified in *S.pyogenes* <SEQ ID 7935> which encodes the amino acid sequence <SEQ ID 7936>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL     Likelihood =-10.99   Transmembrane  450 - 466 ( 445 - 473)
   INTEGRAL     Likelihood = -9.61   Transmembrane   33 -  49 (  30 -  55)
   INTEGRAL     Likelihood = -8.55   Transmembrane  326 - 342 ( 321 - 346)
   INTEGRAL     Likelihood = -7.01   Transmembrane  288 - 304 ( 286 - 311)
   INTEGRAL     Likelihood = -6.79   Transmembrane   95 - 111 (  88 - 114)
   INTEGRAL     Likelihood = -4.99   Transmembrane  265 - 281 ( 264 - 285)
   INTEGRAL     Likelihood = -4.62   Transmembrane  208 - 224 ( 204 - 228)
   INTEGRAL     Likelihood = -3.13   Transmembrane  126 - 142 ( 126 - 145)
   INTEGRAL     Likelihood = -2.81   Transmembrane  366 - 382 ( 365 - 383)
   INTEGRAL     Likelihood = -2.34   Transmembrane  419 - 435 ( 417 - 438)

----- Final Results -----
                bacterial membrane --- Certainty=0.5394(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

A related sequence was also identified in GAS <SEQ ID 9169> which encodes the amino acid sequence <SEQ ID 9170>. Analysis of this protein sequence reveals the following:

```
Possible cleavage site: 39
>>> Seems to have a cleavable N-term signal seq.
     INTEGRAL     Likelihood =-10.99   Transmembrane  443 - 459 ( 438 - 466)
```

```
INTEGRAL    Likelihood = -8.55   Transmembrane  319 - 335 ( 314 - 339)
INTEGRAL    Likelihood = -7.01   Transmembrane  281 - 297 ( 279 - 304)
INTEGRAL    Likelihood = -6.79   Transmembrane   88 - 104 (  81 - 107)
INTEGRAL    Likelihood = -6.99   Transmembrane  258 - 274 ( 257 - 278)
INTEGRAL    Likelihood = -6.62   Transmembrane  201 - 217 ( 197 - 221)
INTEGRAL    Likelihood = -3.13   Transmembrane  119 - 135 ( 119 - 138)
INTEGRAL    Likelihood = -2.81   Transmembrane  359 - 375 ( 358 - 376)
INTEGRAL    Likelihood = -2.34   Transmembrane  412 - 428 ( 410 - 431)

----- Final Results -----
            bacterial membrane --- Certainty=0.539(Affirmative) < succ>
              bacterial cutside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAG08853 GB:AE004959 probable citrate transporter [Pseudomonas aeruginosa]
 Identities = 199/468 (42%), Positives = 296/468 (62%), Gaps = 41/468 (8%)

Query: 9   LLTMLAYAMIIVFMVVMKKMTPFTALVMIPLIMTIAVILTGSADFNADAKFVAFVGDG 68
           +LT+LA+AM+  FM++++N K+++    AL+++P                +AF   G
Sbjct: 1   MLTLLAFAMVATTMFLIMTKRLSALIALILVP----------------------IAFALIG 39

Query: 69  GIAEDLTAIGPWVMYGIINTTAKTGIMLLFAILPFSVMLDAGLPDPITEKMIRFAKGDPMK 128
           G A  L   GPM++ GI   A TG+ML+FAILA+F++M+D+GLPDP   K++R  KGDP+K
Sbjct: 40  GFAAGL---GPMMLDGIRTLAPTGVMLMFAILYFAIMIDSGLPDPAVRKILRLVKGDPLK 96

Query: 129 VLIATAVVAAAVSLNGDGTTTTLICCSAFLPIYKKLDMKIMNLGVLIILQNTIMNLLPWG 188
           V + TA +A  VSL+GDG+TT +IC +A LP+Y +L M  + +  LI+L + ++N+ PWG
Sbjct: 97  VSLGTAALAMIVSLDGDGSTTYMICVAAVLPLYSRLGMSPLVMACLIMLSSGVLNMTPWG 156

Query: 189 GPTARAMSVLGVGP-EILGYLAPGMILSLL--YVICWVAPSMGRKERARLGVIDL--SEE 243
           GPTARA S L V P +I    + P MI   LL   + I W+      G++RRARLG + L    E
Sbjct: 157 GPTARAASALSVDPADIFVPMIPAMIAGLLAIPAIAWI---YGKRSRARLGELHLPTDHE 213

Query: 244 DMRQLTDITDPDTLFIRRPKNFVFNAILTIGLITWLVAGSFNKSIAMAFLLLFAVGTCIA 303
           D+ +++     P+     RRPK   FNAILT+ L+   L+AG      + M  L + A G  IA
Sbjct: 214 DLAEISVSQYPEA---RRPKLLWFNAILTVVLMATLIAGL----LPMPVLFMIAFG--IA 264

Query: 304 LMVNYPVLKDQSKRIGDNAGDAVQVVILVFAAGIFMGLPQGSGMASALAQSFATIIPKQL 363
           ++VNYP +++Q KRIG +A + + VV L+FAAG+F G+  G+GM  A+++S   +IP  L
Sbjct: 265 MIVNYPCIQSQKKRIGAHAENILAVVSLIFAAGVFTGILSGTGMVDAMSKSLLAVIPPAL 324

Query: 364 AGFWGLVIALVSARGTFFISNDGFYYGILPVLAEAGAEYGFSNMAMALASLMGQAFHLLS 423
           +    + ALVS P TFF+SND FYYG+LP+L +A AEYG + + MA AS++GQ  HLLS
Sbjct: 325 GPYLATTTALVSMPFTFFMSNDAPYYGVLPILTQAAAEYGITPVEMARASIVGQPVHLLS 384

Query: 424 PLVAFIYLLLRLTGLDWGEWQKEAAKVALIIFVIPVVTIIAMGQMPLV 471
           PLV   YL+ L  +D G+ Q+   K+A+++ +   + +G  PL+
Sbjct: 385 PLVPSTYLLVGLAKIDPGDHQRFTLKWAVLVCLAILAMALLGLFPLF 432
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2719

A DNA sequence (GASx1174) was identified in *S.pyogenes* <SEQ ID 7937> which encodes the amino acid sequence <SEQ ID 7938>. Analysis of this protein sequence reveals the following:

```
Possible site: 57

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
```

```
bacterial cytoplasm --- Certainty=0.3948(Affirmative) < succ>
 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2720

A DNA sequence (GASx1175) was identified in *S.pyogenes* <SEQ ID 7939> which encodes the amino acid sequence <SEQ ID 7940>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.3519(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2721

A DNA sequence (GASx1177) was identified in *S.pyogenes* <SEQ ID 7941> which encodes the amino acid sequence <SEQ ID 7942>. Analysis of this protein sequence reveals the following:

```
Possible site: 60

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -9.24   Transmembrane  115 - 131 ( 105 - 137)
   INTEGRAL    Likelihood = -8.92   Transmembrane  208 - 224 ( 204 - 238)
   INTEGRAL    Likelihood = -7.80   Transmembrane  282 - 298 ( 273 - 303)
   INTEGRAL    Likelihood = -4.94   Transmembrane   85 - 101 (  75 - 102)
   INTEGRAL    Likelihood = -4.04   Transmembrane   10 -  26 (   3 -  32)
   INTEGRAL    Likelihood = -3.61   Transmembrane  255 - 271 ( 253 - 271)

----- Final Results -----
               bacterial membrane --- Certainty=0.4694(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB89172 GB:AE000960 oxaloacetate decarboxylase, sodium ion pump
             subunit (oadB) [Archaeoglobus fulgidus]
 Identities = 190/354 (53%), Positives = 255/354 (71%), Gaps = 8/354 (2%)

Query: 16  IVMMVIGALLMYLGIKKEYEPTLLVPMGLGTTLVNFPGSGVLMQVVNGVEQEGVFEALFN 75
           +VM+ +G LL+YLGI K+ EP LLVP+G+G ILVN PG G+         E+  +F+
Sbjct: 5   LVMIGVGLLLVYLGIVKKMEPLLLVPIGIGAILVNIPGGGL-------AEEGSIFDLFLK 57
```

```
Query: 76   FGIGTKLFPLLIPIGIGAMIDPGPLLQNPPMLLFGDAAQFGIFVVVVAVLAGFDIKEAA 135
            + I TE+ PLLIP+G+GA+ DF PLL NP   L G AAQ GIF  ++ A+  GF  +EAA
Sbjct: 58   YLIHTEIVPLLIPLGLGALTDFSPLLANPKTPLLGAAAQIGIFAALIAALFLGFTPQEAA 117

Query: 136  SIGIIGAADGPTSIPVANQLAKDNLGPITVAAYSYMALVPIIQPFAIKLVTTKKERIRM 195
             SIGIIG ADGPT+I+    LA  LL    VAAYSYM+LVPIIQP  IK +T+ +ER+I+M
Sbjct: 118  SIGIIGGADGPTTIYTTTILAPHLLAATAVAAYSYMSLVPIIQPPIIKALTSSRERKIKM 177

Query: 196  TYKAENVSQMTKILFPIITLVAGPIAPISLPLVGPLMPGNLLRECGVLDRLSQTAQNEL 255
               +   VS+  KILFPI   +++GP+AP +LPLVG LM GNL RE GV DRL++ A  EL
Sbjct: 178  R-QLRIVSKKKKILFPIATIISGPLAPKALPLVGMLMTGNLPRESGVTDRLAKGASEEL 236

Query: 256  VNIISILLGLTISIKMQADLFLNVQTLLIIVPGLLAFIMDSIGGVMFAKPLNLFRKEKIN 315
            +NI++I+LGL++   M+A+ FL  +TLL++  G++AF   + GGV+ AK +NLF KEKIN
Sbjct: 237 MNIMTIILGLSVGSTMRARSFLTQKTLLVLALGVVAFAAATAGGVLLAKVMNLFLKEKIN 296

Query: 316  PMIGAAGISAFMSSRVIQKMATDEDPQNFILMVAVGANVSGQLASVIAGGLLL 369
             PMIGAAG+SA FMS+RV+Q++A +EDP N ILM+A+G NV+G I S +A G+L+
Sbjct: 297  PMIGAAGVSAVPMSARVVQRLAIEEDPHNHILMHAMGFNVAGVIGSAVAAGVLI 350
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2722

A DNA sequence (GASx1178) was identified in *S.pyogenes* <SEQ ID 7943> which encodes the amino acid sequence <SEQ ID 7944>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -9.50   Transmembrane   21 -  37 (   8 -  43)

----- Final Results -----
               bacterial membrane --- Certainty=0.4800(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2723

A DNA sequence (GASx1179) was identified in *S.pyogenes* <SEQ ID 7945> which encodes the amino acid sequence <SEQ ID 7946>. Analysis of this protein sequence reveals the following:

```
Possible site: 60

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1906(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF93961 GB:AE004165 citrate lyase, gamma subunit [Vibrio cholerae]
 Identities = 46/97 (47%), Positives = 64/97 (65%)

Query: 1  MDIKQTAVAGSLESSDLMQTVSPNDEQTITFTLDSGVEKQFGNHIRQLLHQTLVNLKVTA 60
          M I    A AG+LESSDL + + PN++  I + LDS+VR+QFG+ IRQ++  TL  ++V
Sbjct: 1  MKLAHPAFAGTLESSDLQVRIEPNNDGGIELVLDSTVEQQFGHAIRQVVLHTLDAMQVRD 60

Query: 61 AKVEAVDKGALDCTIQARTIAAVHRAAGIDQYDWKEI 97
           A V   DKGALDC I+AR  AAV RA  +   +W ++
Sbjct: 61 ALVTIEDKGALDCVIRARVQAAVMRACIVQNIEWSQL 97
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2724

A DNA sequence (GASx1181) was identified in *S.pyogenes* <SEQ ID 7947> which encodes the amino acid sequence <SEQ ID 7948>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.65   Transmembrane   74 -  90 (  74 -  90)

----- Final Results -----
               bacterial membrane --- Certainty=0.1659(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA71632 GB:Y10521 CILB, citryl-CoA lyase beta subunit
          [Leuconostoc mesenteroides]
 Identities = 187/293 (63%), Positives = 237/293 (80%), Gaps = 1/293 (0%)

Query: 2   ERLRRTMMPVPGANAAMLRDAPLPGADSIMFDLEDSVSLKEKDTSRALVHPALKTFDYSS 61
           ERLRRTMMPVPG N AM++DA +PGADSIMFDLED+VSL EKD++R LV+ AL+T DY S
Sbjct: 4   ERLRRTMMPVPGNNPAMVKDAGIPGADSIMFDLEDAVSLAEKDSARYLVYEALQTVDYGS 63

Query: 62  VETVVRVNGLDS-CGALDIEAVVLAGVNVIRLPKETAQDIIDVEAVIERVERENSIEVG 120
            E VVR+NGLD+     DI+A+V AG++VIRLPK ETA  + ++E++I   E+E    VG
Sbjct: 64  SELVVRINGLDTPFVKNDIKAMVKAGIDVIRLPKVETAAMMHELRSLITDAEKEPGRPVG 123

Query: 121 RTRMMAAIESABGVLNARETAKASKRLIGIALGAEDVVTNMKTRYPDGQELFFARSMIL 180
            T MMAAIESA GV+NA EIA AS R+IGIAL AEDV T+MKT RYPDGQEL +AR++IL
Sbjct: 124 TTHMMAAIESALGVVNAVEIANASDRMIGIALSAEDYTTDMKTHRYPDGQELLYARNVIL 183

Query: 181 HAARAAGIAAIDTVYSDVNNTEGFQNEVRMIKQLGFDGKSVINPRQIPLVNEIYTPTKKE 240
           HAARAAGIAA DTV+++N+ EGF  E ++I QLGFDGKS+INPRQI +VN++Y PT+KE
Sbjct: 184 HAARAAGIAAPDTVFTNLNDEEGFVRETQLIHQLGFDGKSLINPRQIEMVNKVYAPTEKE 243

Query: 241 IDHAKQVIWAIRRAESKGSGVISLNGKMVDKPIVERAKRVIALATAAGVLSEE 293
           I++A+ VI AI EA+ KGSGVIS+NGKMVD+P+V RA+RV+ LA A  ++  E
Sbjct: 244 INNAQNVIAAIEEAKQKGSGVISMNGQMVDRPVVLRAQRVMKLANANHLVDSE 296
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2725

A DNA sequence (GASx1182) was identified in *S.pyogenes* <SEQ ID 7949> which encodes the amino acid sequence <SEQ ID 7950>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3554(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA71633 GB:Y10621 CILA, citrate CoA-transferase alpha subunit
          [Leuconostoc mesenteroides]
 Identities = 294/511 (57%), Positives = 378/511 (73%), Gaps = 7/511 (1%)

Query: 4   NKLGRDIPQPVADQY--GVPEGELANIKQYDESSRRIKFVKPGDSKLLGSVREAIEKTGL 61
           NK+  D+P    +Q    VPE        +       +++     G+SK+  S+ + +  T L
Sbjct: 3   NKVNIDVPDALLSQLDDSVPESTNYGNPEIQRVGPKVRATT-GESKVQSSIDVLSNT-L 60

Query: 62  TDGMTISFHHHFRBGDFIMNMVLEEIAKMGIKNLSIAPSSIANV-HEPLIDHIKNGVVTN 120
            DGMTISFHHHFRBGDF+ N V+ +I  MG +NL++APSS+ NV ++ +I+ IK GVVTN
Sbjct: 61  KDGMTISFHHHFRBGDFVFNKVMRKIIDMGYQNLTLAPSSLTNVMNDIVIEAIKKGVVTN 120

Query: 121 ITSSGLRDKVGPAISEGLMENPVVIRSHGGRARAIASGDIHIDVAFLGAPSSDAVGNVNG 180
            ITSSG+R  +G A+S G++NPV+ RSHG RARAI SG+I IDVAFLG P+SD  GN NG
Sbjct: 121 ITSSGMRGTLGDAVSHGILKNPVIFRSHGARARAIESGEIKIDVAFLGVPNSDEMGNANG 180

Query: 181 TKGKATCGSLGYAMIDAKYADQVVILTDNLVPYPNTPISIPQTDVDVVVTVDAIGDPQGI 240
             G A  GSLGYA+IDA+YAD++V++TD ++PYPNTP SI QT VDYVV VD +GDP  I
Sbjct: 181 MNGDAAFGSLGYALIDAQYADKLVLITDTIMPYPNTPASIKQTQVDYVVKVDKVGDPDKI 240

Query: 241 AKGATRFTKNPKELLIAEYAAKVITNSPYFKEGFSPQTGTGGASLAVTRFMREAMIKENI 300
             GATRFTK+PKEL IA+    VI NS YFK  FSPQTG+GGA+LAVTRF+REAM+ +NI
Sbjct: 241 GSGATRFTKDPKELKIAKTVNDVIVNSKYFKNDFSPQTGSGGAALAVTRFLREAMVAQNI 300

Query: 301 KASFALGGITNAMVELLSEELVEKILDVQDFDHPSAVSLGKHARHYEIDANMYASFLSKG 360
             ASFALGGIT   V+LL E LV ++DVQDFD  +A S+       EIDA+ YA P +KG
Sbjct: 301 MASFALGGITKPTVDLLNEGLVNRVMDVQDFDKGAASSMKLSPNQQEIDASWYADPANKG 360

Query: 361 AVINQLDTCILSALEVDTNFNVNVMTGSDGVIRGASGGHCDTAFAAKMSLVISPLIRGRI 420
            A++++LD  ILSALEVDTNFNVNM+GSDGVIRGA GGH D A  AK+++++  PL+RGRI
Sbjct: 361 AMVDKLDVAILSALEVDTNFNVNMVGSGSDGVIRGAIGGHQDAA-TAKLTIISVPLVRGRI 419

Query: 421 PTFVDEVNTVITPGTSVDVIVTEVGIAINPNRQDLVDHPFKSL-NVPQFSIEKLKEKAYAI 479
             T V +VNTVITPG S+DV+VTEVGIAINP R DLV+  K +  +P +SIEEL++KA  I
Sbjct: 420 ATIVPKVNTVITPGDSIDVVVTEVGIAINPKRTDLVEQLKQVPGLPIYSIEELQQKAEKI 479

Query: 480 VGTPERIQYGDKVVALIEYRDGSLMDVVYNV 510
            VG P  +++ D+VVA+ EYRDGS+++D++  V
Sbjct: 480 VGQPAPLKPTDRVVAAEYRDGSVIDIIKEV 510
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2726

A DNA sequence (GASx1183) was identified in *S.pyogenes* <SEQ ID 7951> which encodes the amino acid sequence <SEQ ID 7952>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA71634 GB:Y10621 CILG, hypothetical protein [Leuconostoc
          mesenteroides]
 Identities = 65/176 (36%), Positives = 97/176 (54%), Gaps = 3/176 (1%)

Query: 21  DTYPSGEAIQLSDMLRAREERALRQLHLLKEYPEGSLLSVTMNIPGPIKTSPKLLEAFDI 80
           D +  GE + L  +L  RE R   Q  L+  +P   + SV +N+PGPIKTSPKL   F I
Sbjct: 2   DYPEGGERLNLMQVLDNREWREKYQKQLMASFPTAVITSVKLNLPGPIKTSPKLQSVFQI 61

Query: 81  VIKAIQTALADDKICYQLRLL-PTGYEYYLITSLPSRDLKLKMIALETELPIGRLMDLD 139
           +I  +      D +I  +   +   TG + + +TS   + +K  MI  E     +GRL+DLD
Sbjct: 62  IINDLNPVFKDLQIIKEASFVDQITGPDIPFVTSGCLKLVKQIMITFEESHLLGRLLDLD 121

Query: 140 VLVLQNDLPHSISRTVLGGSPRQCFICSKEAKVCGRLRKHSVEEMQTAISKLLHSP 195
           V+    D    +SR  LG +PR+C +C K+AK C +   HS+ E  + I+K+LH+P
Sbjct: 122 VMCQNAD--KQLSREELGFAPRKCLLCGKDAKTCIKEGNHSLAEGYSQINKMLHNF 175
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2727

A DNA sequence (GASx1184) was identified in *S.pyogenes* <SEQ ID 7953> which encodes the amino acid sequence <SEQ ID 7954>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.3730(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB99233 GB:U67563 oxaloacetate decarboxylase alpha chain (oadA)
          [Methanococcus jannaschii]
 Identities = 245/441 (55%), Positives = 336/441 (75%), Gaps = 5/441 (1%)

Query: 10  IRITETVLRDGQQSQIATRMTTKEMIPILETLDNAGYHALEMWGGATFDSCLRFLNEDPW 69
           ++I +T  RD QQS IATRM T++M+PI E +D  G++++E+WGGATFD+C+R+LNEDPW
Sbjct: 2   VKIVDTTFRDAQQSLIATRMRTEDMLPIAEKMDEVGFYSMEVWGGATFDACIRYLNEDPW 61

Query: 70  ERLRAIRKAVKKTKLQMLLRGQNLLGYRNYADDVVRSFIQKSIRNGIDIVRIFDALNDPR 129
           ERLRA++K ++ T LQMLLRGQNL+GYR+Y DD+V  F+ K+ ENGIDI RIFDALND R
Sbjct: 62  ERLRALKKRIQNTPLQMLLRGQNLVGYRHYPDDIVEKFVIKAHENGIDIFRIFDALNDVR 121

Query: 130 NLQTAVSATKKPGGHAQVAISYTTSPVHTIDYFVELAKAYQAIGADSICIKDMAGVLTPE 189
           N++TA+   KK G   Q AI YT SPVHTID +VELAK  + +G DSICIKDMAG+LTP
Sbjct: 122 NMETAIKTAKKVGAEVQGAICYTISPVHTIDQYVELAKKLEEMGCDSICIKDMAGLLTPY 181
```

```
Query: 190 IGYQLVKCIKENTTIPLEVHTHATSGISEMTYLKVAEAGADIIDTAISSPSGGTSQPATE 249
           GY+LVK +KE   ++P++VH+H TSG++ MTYLKV EAGAD++D AIS P+ GTSQP TE
Sbjct: 182 EGYELVKRLKSEISLPIDVHSHCTSGLAPMTYLKVIEAGADMVDCAISPFAMGTSQPPTE 241

Query: 250 SMAIAIJTDLGFDTGLDMQSVAKVAEYFNTIRDHYRKIGILNPKVKLDTEPKTLIYQVPGGM 309
           S+ +AL     +DTGLD++ + ++ +YF  +R+ Y+    + +P  + + + L+YQVPGGM
Sbjct: 242 SIVVALKGTKYDTGLDLKLLNEIRDYFMRVRKKYKM--LPSPISQIVDARVLVYQVPGGM 299

Query: 310 LSNLLSQLTEQGLTDKYEEVLAEVPKVRADLGYPPLVTPLSQMVGTQAVLMNVLSGERYKV 369
           LSNL+SQL EQG  DK+EEVL E+P+VR DLGYPPLVTP SQ+VGTQA++N+++ ERYK+
Sbjct: 300 LSNLVSQLKEQGSALDKFEEVLQEIPRVRKDLGYPPLVTPTSQIVGTQAVLNVLTERRYKI 359

Query: 370 VPNEIKDYVRGLYGQSPAPLAEGIKEKIIGD-EAVITCRPADLIEPQMIYLRDEIAP--Y 426
           + NE+ +YV+G YG+ PAP+   + ++++ + E  ITCRPADL+ P+    ++ E
Sbjct: 360 ITNEVVNYVKGFYGKPPAPINPELLKRVLDEGEKPITCRPADLLPPEWEKVKKEAEEKGI 419

Query: 427 AHSEEDVLSYASFPQQARDFL 447
             EED+L+YA +PQ A  FL
Sbjct: 420 VKKEEDILTYALYPQIAVKFL 440
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2728

A DNA sequence (GASx1185R) was identified in *S.pyogenes* <SEQ ID 7955> which encodes the amino acid sequence <SEQ ID 7956>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2497(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF93960 GB:AE004165 citrate (pro-3S)-lyase ligase [Vibrio cholerae]
 Identities = 118/336 (35%), Positives = 183/336 (54%), Gaps = 5/336 (1%)

Query: 4   YTISKVFPSDKTTMASVKNLLKQEGIRLDARHLDYTCAIMNAQNDVIATGSYPGNSLRCLC 63
            YT S+V   ++T +  +K  L Q  + +D  +++    +    N +IA G   G+ L+ +
Sbjct: 10  YTFSRVSTCNRTKLLQIKEFLCQHQLTVDDDVEHF-VVAYGTNQIIACGGIAGHVLKSIA 68

Query: 64  VSSAYQGEGLLNRIVSHLIDEEYALGNYHLFVYTKTSSAAFFKDLGFTEIVHIDNHISFL 123
            VS A QC G   ++++ L +  Y +G + LF++TK ++    F+  GF  +  ++ HI+ L
Sbjct: 69  VSPALQGTGFALKLMTELTNPAYEMGRFSLFLFTKPANIDLFRQCGFFLVDKVEPHIALL 128

Query: 124 ENKKTGFQDYLMTLNKPEQTPGKVAAIVINANPFTLGHQFLVEKAARENDWVHLFMVSED 183
            EN      Y   L   +  +  K+ +IV+NANPFTLGHQ+L+E+A   + DWVHLF+V  +
Sbjct: 129 ENSPNRLSVYCKQLQLLKMSGRKIGSIVMNANPFTLGHQYLIEQACEQCDWVHLFVVKAE 188

Query: 184 RSLIPFSVRKRLIQEGLAHLDNVIYHETGPYLISQATFPAYFQKDDNDVIKSQALLDTNI 243
                ++ R  +I+ G  HL N+  H     Y+LS+ATFP+YF K+   V +S   LD +I
Sbjct: 189 NKDFSYADRMAMIKAGSKHLLNLTIHSGSDYIISRATFPSYFIKDQQVVNQSHTALDLSI 248

Query: 244 FL-KIAQTLQITKRYVGEEPTSRVTAIYNEIM---AEQLQQAGILLDILPRKAINQQDP 299
            F   IA  L IT R+VG EP   VT  YN+ M    E+   A  + ++  +  Q  P
Sbjct: 249 FRHSIAPALGITHRFVGSEPICTVTRHYNQAMRRWLEEAHDASAPIQVVEIERSQQASQP 308

Query: 300 ISASTARQALKDNDWDLLAKLLPKTSLDYFCSLKAQ 335
```

```
          ISAS  R  LK   +  +A L+PKT+  Y C   A+
Sbjct: 309 ISASRVRYLLKQPGFAAIADLVPKTTYSYLCQHYAE 344
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2729

A DNA sequence (GASx1187) was identified in *S.pyogenes* <SEQ ID 7957> which encodes the amino acid sequence <SEQ ID 7958>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4790(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2730

A DNA sequence (GASx1188R) was identified in *S.pyogenes* <SEQ ID 7959> which encodes the amino acid sequence <SEQ ID 7960>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3956(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2731

A DNA sequence (GASx1190) was identified in *S.pyogenes* <SEQ ID 7961> which encodes the amino acid sequence <SEQ ID 7962>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1274(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2732

A DNA sequence (GASx1196R) was identified in *S.pyogenes* <SEQ ID 7963> which encodes the amino acid sequence <SEQ ID 7964>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2733

A DNA sequence (GASx1211) was identified in *S.pyogenes* <SEQ ID 7965> which encodes the amino acid sequence <SEQ ID 7966>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial cytoplasm --- Certainty=0.1850(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2734

A DNA sequence (GASx1219R) was identified in *S.pyogenes* <SEQ ID 7967> which encodes the amino acid sequence <SEQ ID 7968>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial cytoplasm --- Certainty=0.2284(Affirmative) < succ>
```

```
bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2735

A DNA sequence (GASx1225) was identified in *S.pyogenes* <SEQ ID 7969> which encodes the amino acid sequence <SEQ ID 7970>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2062(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2736

A DNA sequence (GASx1229) was identified in *S.pyogenes* <SEQ ID 7971> which encodes the amino acid sequence <SEQ ID 7972>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2755(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2737

A DNA sequence (GASx1247R) was identified in *S.pyogenes* <SEQ ID 7973> which encodes the amino acid sequence <SEQ ID 7974>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -6.32   Transmembrane   55 -  71 (  53 -  81)
```

```
   INTEGRAL    Likelihood = -6.00   Transmembrane   74 -  90 (  72 -  95)
   INTEGRAL    Likelihood = -2.18   Transmembrane   95 - 111 (  95 - 111)
   INTEGRAL    Likelihood = -1.54   Transmembrane  124 - 140 ( 123 - 141)

----- Final Results -----
               bacterial membrane --- Certainty=0.3527(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB14326 GB:Z99116 yqjA [Bacillus subtilis]
 Identities = 97/306 (31%), Positives = 154/306 (49%)

Query: 6   RTLKMTLATIVAILIAYQLHLDYAMSAGIIALLSVLDTRKSSLVVARNKLLSFFLAFGIA 65
           RT+K  L T +AI I+  LHL    SAGII +L +  T+K SL  +  R  +  LA   +
Sbjct: 7   RTIKTALGTALAIYISQLLHLQNFASAGIITLLCIQITQKRSLQASWARFWACCLAIAFS 66

Query: 66  MMCFSLFGFTTVGFMCYLLIIPLLYHFQIEAGLVPITVLVTHLIAKKSIALPILSNEFM 125
            + F L G+        LLI IP+    +I  G+V  +V++ HL      I    + NE
Sbjct: 67  YLFFELIGYHPFVIGALLLIPIPITVLLKINEGIVTSSVIILHLYMSGGITPTFIWNEVQ 126

Query: 126 LFFVGTSVALLFNAYMGPQDQQIRYHQKVESDLKGILYRFESFLLEGKQNEGLLIKNL 185
           L  VG   VALL N YM   D+++  Y +K+E +   I    E +LL G+   G  I
Sbjct: 127 LITVGIGVALLMNLYMPSLDRKLIAYRKKIEDNFAVIFAEIERYLLITGEQDWSGKEIPST 186

Query: 186 DKILDEALKLVYRERHNQLPQQTNYCVHYFEMRRQQNRLLGQMAINVNTLMRQSKESSIL 245
            +++ EA  L YR+  N + +  N   HYF+MR +Q  ++ ++   V ++     +  ++
Sbjct: 187 HQLITEAKNLAYRDVQNHILRYENLHYHYFKMREKQFEIIERLLPKVTSISITVDQGKMI 246

Query: 246 SHLPHETACQLSEQNPALTLIDDIEQLLETFRHGDLPQTREEFERRAVLFQLLQDLERFI 305
           +   H+    +  N A   +  +  + + F    LP TREEFE RA LF LL ++E+++
Sbjct: 247 AEFIHDLREAIHPGNTAYKFLKRLADMRKEFEEMPLPATREEFSARAALFHLLGEMEQYL 306

Query: 306 LLKVEF 311
           ++K  F
Sbjct: 307 VIKSYF 312
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2738

A DNA sequence (GASx1261) was identified in *S.pyogenes* <SEQ ID 7975> which encodes the amino acid sequence <SEQ ID 7976>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.6082(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2739

A DNA sequence (GASx1262R) was identified in *S.pyogenes* <SEQ ID 7977> which encodes the amino acid sequence <SEQ ID 7978>. Analysis of this protein sequence reveals the following:

```
Possible site: 51

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -7.06   Transmembrane   38 -  54 (  37 -  55)

----- Final Results -----
                bacterial membrane --- Certainty=0.3824(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2740

A DNA sequence (GASx1265R) was identified in *S.pyogenes* <SEQ ID 7979> which encodes the amino acid sequence <SEQ ID 7980>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                 bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2741

A DNA sequence (GASx1270) was identified in *S.pyogenes* <SEQ ID 7981> which encodes the amino acid sequence <SEQ ID 7982>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.4063(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2742

A DNA sequence (GASx1290R) was identified in *S.pyogenes* <SEQ ID 7983> which encodes the amino acid sequence <SEQ ID 7984>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-12.37   Transmembrane  180 - 196 ( 172 - 207)
   INTEGRAL    Likelihood =-10.19   Transmembrane   34 -  50 (  30 -  53)
   INTEGRAL    Likelihood = -4.09   Transmembrane  233 - 249 ( 232 - 250)

----- Final Results -----
              bacterial membrane --- Certainty=0.5946(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB88010 GB:L21856 MalA [Streptococcus pneumoniae]
 Identities = 66/237 (27%), Positives = 105/237 (43%), Gaps = 28/237 (11%)

Query: 45  MIPVTLHYANMTTYPLERIVTKSLSPITDKTYQALTQGKIEKD---TPQQQSLIRRD--- 98
           M+P+ +  ++  TYPLE  +      P+TDK  Q L++      D   T+ G +
Sbjct: 1   MVPIAIQNSSQETYPLETFIDNVYEPLTDKVVQDLSEHATIVDGTLTYTGTASQAPSVVI 60

Query: 99  GELVLAVLPTKVDLEQLASESTRQIIVTKKEWRFVTPDGKEL-RAHVRGQQQSLADLTTV 157
           G   +  LP  + L        T ++++++K        + KEL R   R  Q       T
Sbjct: 61  GPSQIKELPKDLQLHF----DTNSLVISK--------ESKELTRISYRAIQ------TES 102

Query: 158 KAVKDFVNQQNY---DSNKASVLGFLLLTFVLMVCVGTLIVIGLGAFFLTLTKRSRLFMI 214
               KD + Q +      +N+  +  FL+L     +  +    IV        L +TK+SRLF
Sbjct: 103 FKSKDSLTQAPIRLVPTNRVYISLFLVLGASFLFGLNFFIVSLGACLLLYITKKSRLFSF 162

Query: 215 RNFSEGLGLMVNCLAWPSLLAIALSFFIQDPVLIMNCQVFGTLLMLTWVFYKTQFRD 271
           R F E     ++NCL  P+L+ + L   F Q+   ++  Q    +L L  +FYKT FRD
Sbjct: 163 RTFKECYHFILNCLGLPTLITLILGLFGQNMTTLTTVQNILFVLLVTIPYKTHFRD 219
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2743

A DNA sequence (GASx1294) was identified in *S.pyogenes* <SEQ ID 7985> which encodes the amino acid sequence <SEQ ID 7986>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial cytoplasm --- Certainty=0.2104(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2744

A DNA sequence (GASx1303R) was identified in *S.pyogenes* <SEQ ID 7987> which encodes the amino acid sequence <SEQ ID 7988>. Analysis of this protein sequence reveals the following:

```
Possible site: 38

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -8.07   Transmembrane   13 -  29 (   8 -  38)

----- Final Results -----
               bacterial membrane --- Certainty=0.4227(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2745

A DNA sequence (GASx1307R) was identified in *S.pyogenes* <SEQ ID 7989> which encodes the amino acid sequence <SEQ ID 7990>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2746

A DNA sequence (GASx1312R) was identified in *S.pyogenes* <SEQ ID 7991> which encodes the amino acid sequence <SEQ ID 7992>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1996(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2747

A DNA sequence (GASx1316R) was identified in *S.pyogenes* <SEQ ID 7993> which encodes the amino acid sequence <SEQ ID 7994>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3504(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>

RGD motif: 271-273
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC66321 GB:AE000792 outer surface protein, putative [Borrelia
          burgdorferi]
 Identities = 127/365 (34%), Positives = 195/365 (52%), Gaps = 14/365 (3%)

Query: 1   MVDLGFSLYPERYDVTKSKAYIDLCHSYGAKRLFMSLLQLAPADHQMFHCYAELIAYANQ 60
            M ++G S+YP       K   Y++     +G  ++F SLL +   +   F  + EL++ AN+
Sbjct: 1   MKEIGISIYPMVSPKNKIKYLEKSAHRGFTQVFTSLLYI---NGNEFDIFKELLSIANK 57

Query: 61  LGIRVIADVSPSFISQAGWSDQLIERA------HAFGLAGLRLDEALPAEIVTLTNPFF 114
             G++ I DVSP    + G    +              G   +RLD      E   +T N
Sbjct: 58  NGMKPIIDVSPEIFKELGIDLSNLRNCPKLDYFKKLGAWAIRLDNTFTGIEESLNTFNDS 117

Query: 115 GLKIELNMSTDKQLLMSLLATDAERSNIIGCHNFYPHEFTGLSWQHFKDMGRFYHEHDIE 174
             LKI+LN+S   + + +++        N++GCHNFYPH++TGLS   FK+ ++ +  + I
Sbjct: 118 DLKIQLNISNINKHIDTIMYFKPNIKNLLGCHNFYPHKYTGLSRNFFKETTKIFKHYSIP 177

Query: 175 TAAFITAQSASE-GPWLLAEGLPTVEDHRHLPIGLQVELMKAIGTIDNILISNQPISEEE 233
            TAAFI++ +A E       EG+PT+E HR   I  Q + +   G ID +LISN F SE E
Sbjct: 178 TAAFISSNNAEECARGKEKEGVPTLESHRSKDIETQAKDLFKEG-IDTVLISNCFPSETE 236

Query: 234 LAACTQALARPVTTIKVRPIIDLTEVEBQII-GYPHCYRGDVSDYVIRSTMPRLVYAQES 292
            L   ++ + R +  +K     D   VE++II    H  RGD++ Y IRSTMPR+ Y  +
Sbjct: 237 LKKVSK-VNRNILELKADLNPDANSVEKEIILENLHFNRGDINSYRIRSTMPRVYTYNKK 295

Query: 293 IAPRDQSKEVKRGSIIIDNDRVHRYKGRLQIALKNFTVSSKANVVARVRERDYLSLLDDLR 352
             P     E+K+G I+ID+  Y  Y GELQIALK+   +   NVV ++  D + LL+ +
Sbjct: 296 F-PVHSPNEIKKGDILIDSSEYLGYTGRLQIALKDTPNNGLVNVVGKIINDEIYLLEKIE 354

Query: 353 FWQEF 357
            FW++F
Sbjct: 355 FWEKF 359
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2748

A DNA sequence (GASx1319) was identified in *S.pyogenes* <SEQ ID 7995> which encodes the amino acid sequence <SEQ ID 7996>. Analysis of this protein sequence reveals the following:

```
Possible site: 34
```

```
>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -9.50   Transmembrane  127 - 143 ( 125 - 151)
   INTEGRAL    Likelihood = -7.43   Transmembrane   17 -  33 (  15 -  36)
   INTEGRAL    Likelihood = -5.68   Transmembrane   39 -  55 (  36 -  57)
   INTEGRAL    Likelihood = -1.86   Transmembrane   60 -  76 (  59 -  77)
   INTEGRAL    Likelihood = -0.59   Transmembrane   85 - 101 (  85 - 101)

----- Final Results -----
               bacterial membrane --- Certainty=0.4800(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2749

A DNA sequence (GASx1320) was identified in *S.pyogenes* <SEQ ID 7997> which encodes the amino acid sequence <SEQ ID 7998>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.81   Transmembrane   35 -  51 (  35 -  51)

----- Final Results -----
               bacterial membrane --- Certainty=0.1723(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2750

A DNA sequence (GASx1321) was identified in *S.pyogenes* <SEQ ID 7999> which encodes the amino acid sequence <SEQ ID 8000>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2751

A DNA sequence (GASx1329) was identified in *S.pyogenes* <SEQ ID 8001> which encodes the amino acid sequence <SEQ ID 8002>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.28   Transmembrane   64 -  80 (  64 -  80)

----- Final Results -----
               bacterial membrane --- Certainty=0.1510(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2752

A DNA sequence (GASx1332R) was identified in *S.pyogenes* <SEQ ID 8003> which encodes the amino acid sequence <SEQ ID 8004>. Analysis of this protein sequence reveals the following:

```
Possible site: 37

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2753

A DNA sequence (GASx1333) was identified in *S.pyogenes* <SEQ ID 8005> which encodes the amino acid sequence <SEQ ID 8006>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2754

A DNA sequence (GASx1335R) was identified in *S.pyogenes* <SEQ ID 8007> which encodes the amino acid sequence <SEQ ID 8008>. Analysis of this protein sequence reveals the following:

```
Possible site: 37

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF96047 GB:AE004354 uridine phosphorylase [Vibrio cholerae]
 Identities = 46/167 (27%), Positives = 72/167 (42%), Gaps = 12/167 (7%)

Query: 8   GVKEMISTGTCGVLVP-IAENRPLVPVKALRDEGTSYHYVAPSRYIDIDPKMLRLIEKTL 66
           G K ++  G+ G +   I     ++   A+RDEG S  Y+  +       +++  +++ L
Sbjct: 79  GAKAIVRVGSAGAMQSEIGLGELILVEGAVRDEGGSKAYIGAAYPAYSSPELVVEMQRFL 138

Query: 67  LAQGLAYQEVITWSTDGFYR-ETKEKVAHRQEEGCSVVEMECSALAAVAQLRG-----IL 120
             Q +     I  S D FY  E  E   +G    +ME SAL  V +LRG     +L
Sbjct: 139 AEQSVPIHRGIVRSHDSFYTDESAELCRYWHRNGILAADMETSALLTVGRLRGLQVASVL 198

Query: 121 WGQLLFTADTLADVEVY---DQRNWGADSFSFALHLCLEVLNTLEKD 164
               +L+  D  A V  Y   DQR    +  + A    L  LN L+ D
Sbjct: 199 NNVVLYEQDVQAGVNQYVNADQRMWQSR--TLAARAALHALNALKFD 243
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2755

A DNA sequence (GASx1353) was identified in *S.pyogenes* <SEQ ID 8009> which encodes the amino acid sequence <SEQ ID 8010>. Analysis of this protein sequence reveals the following:

```
Possible site: 42

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -5.79   Transmembrane  241 - 257 ( 234 - 260)
   INTEGRAL    Likelihood = -5.15   Transmembrane   44 -  60 (  43 -  65)
   INTEGRAL    Likelihood = -4.78   Transmembrane   74 -  90 (  72 -  92)

----- Final Results -----
               bacterial membrane --- Certainty=0.3314(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2756

A DNA sequence (GASx1354R) was identified in *S.pyogenes* <SEQ ID 8011> which encodes the amino acid sequence <SEQ ID 8012>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -3.45   Transmembrane   68 -  84 (  65 -  86)

----- Final Results -----
                bacterial membrane --- Certainty=0.2361(Affirmative) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB83831 GB:AL162753 putative integral membrane protein
          [Neisseria meningitidis]
 Identities = 31/72 (43%), Positives = 46/72 (63%), Gaps = 6/72 (8%)

Query: 17  FVIYAPDKRKAIKKKRRISERKLLVITVLPGGF-GALLAAKKYHHKTRGWYFVI----TC 71
           F +Y  DKR+A++ KRRI E +LL +  LPGG+ GA L ++ + HKT K  FV+     T
Sbjct: 38  FALYGIDKRRAVRGKRRIPEHRLL-LPALPGGWAGAYLGSRIPEHKTAKKRFVVLFRLTV 96

Query: 72  YTSILLTLLVTY 83
             ++L TL++ Y
Sbjct: 97  SGNVLATLILIY 108
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2757

A DNA sequence (GASx1363R) was identified in *S.pyogenes* <SEQ ID 8013> which encodes the amino acid sequence <SEQ ID 8014>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2758

A DNA sequence (GASx1367) was identified in *S.pyogenes* <SEQ ID 8015> which encodes the amino acid sequence <SEQ ID 8016>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have an uncleavable N-term signal seq
```

```
----- Final Results -----
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA63508 GB:X92946 hypothetical protein [Lactococcus lactis]
 Identities = 64/96 (66%), Positives = 77/96 (79%)

Query: 1  MPRKTFDKAFKLGAVKLILEHQPVKMVSSTLEIHPNSLYQWIQEYEKYGESAFPGHGSA 60
          M R+ FDK FK SAVKLILEE   VK VS  LE+H NSLY+W+QE E+YGESAFPG+G+A
Sbjct: 1  MARRKFDKQFKNSAVKLILEEGYSVKEVSQELEVHANSLYRWVQEVEEYGESAFPGNGTA 60

Query: 61 LRHAQFKTKKLEKEHKLLQEELALLKKFQVFLKPNR 96
          L +AQ K K LEKE++ LQEEL LLKKF+VFLK ++
Sbjct: 61 LANAQHKILKLEKENRYLQEELELLKKFRVFLKRSK 96
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2759

A DNA sequence (GASx1374R) was identified in *S.pyogenes* <SEQ ID 8017> which encodes the amino acid sequence <SEQ ID 8018>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2585(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2760

A DNA sequence (GASx1382R) was identified in *S.pyogenes* <SEQ ID 8019> which encodes the amino acid sequence <SEQ ID 8020>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -2.39   Transmembrane    3 -  19 (   3 -  19)

----- Final Results -----
                bacterial membrane --- Certainty=0.1956(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2761

A DNA sequence (GASx1391R) was identified in *S.pyogenes* <SEQ ID 8021> which encodes the amino acid sequence <SEQ ID 8022>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> May be a lipoprotein

----- Final Results -----
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2762

A DNA sequence (GASx1404) was identified in *S.pyogenes* <SEQ ID 8023> which encodes the amino acid sequence <SEQ ID 8024>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial cytoplasm --- Certainty=0.3046(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2763

A DNA sequence (GASx1412R) was identified in *S.pyogenes* <SEQ ID 8025> which encodes the amino acid sequence <SEQ ID 8026>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
             bacterial cytoplasm --- Certainty=0.1590(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2764

A DNA sequence (GASx1414R) was identified in *S.pyogenes* <SEQ ID 8027> which encodes the amino acid sequence <SEQ ID 8028>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2816(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2765

A DNA sequence (GASx1416) was identified in *S.pyogenes* <SEQ ID 8029> which encodes the amino acid sequence <SEQ ID 8030>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1744(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2766

A DNA sequence (GASx1417) was identified in *S.pyogenes* <SEQ ID 8031> which encodes the amino acid sequence <SEQ ID 8032>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3771(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2767

A DNA sequence (GASx1419R) was identified in *S.pyogenes* <SEQ ID 8033> which encodes the amino acid sequence <SEQ ID 8034>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-10.93   Transmembrane    4 -  20 (   1 -  25)

----- Final Results -----
                bacterial membrane --- Certainty=0.5373(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2768

A DNA sequence (GASx1423) was identified in *S.pyogenes* <SEQ ID 8035> which encodes the amino acid sequence <SEQ ID 8036>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -8.97   Transmembrane   30 -  46 (  25 -  49)
   INTEGRAL    Likelihood = -7.80   Transmembrane   52 -  68 (  50 -  72)
   INTEGRAL    Likelihood = -6.95   Transmembrane  129 - 145 ( 125 - 146)

----- Final Results -----
                bacterial membrane --- Certainty=0.4588(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2769

A DNA sequence (GASx1426R) was identified in *S.pyogenes* <SEQ ID 8037> which encodes the amino acid sequence <SEQ ID 8038>. Analysis of this protein sequence reveals the following:

```
Possible site: 25
```

```
>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -3.45   Transmembrane   36 -  52 (  36 -  55)

----- Final Results -----
                bacterial membrane --- Certainty=0.2381(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC39287 GB:AF115103 orf87 gp [Streptoococcus thermophilus
          bacteriophage Sfi21]
 Identities = 43/73 (58%), Positives = 61/73 (82%)

Query: 1  MINLKLRLQNKVTLMAILGAIFLLAQQLGIKLPSNIADIANIAVTLLVLLGVVTDPTTKG 60
          MIN KLRLQNK TL+A++ A+FL+ QQ G+ +P+NI +  NT V +LV+LG++TDPTTKG
Sbjct: 8  MINFKLRLQNKATLVALISAVFLMLQQFGLHVPNNIQEGINTLVGILVILGIITDPTTKG 67

Query: 61 LSDSEQALTYHEP 73
          ++DSE+AL+Y +P
Sbjct: 68 IADSERALSYIQP 80
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2770

A DNA sequence (GASx1427R) was identified in *S.pyogenes* <SEQ ID 8039> which encodes the amino acid sequence <SEQ ID 8040>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -3.03   Transmembrane    2 -  18 (   1 -  23)

----- Final Results -----
                bacterial membrane --- Certainty=0.2211(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2771

A DNA sequence (GASx1428R) was identified in *S.pyogenes* <SEQ ID 8041> which encodes the amino acid sequence <SEQ ID 8042>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.1017(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2772

A DNA sequence (GASx1429R) was identified in *S.pyogenes* <SEQ ID 8043> which encodes the amino acid sequence <SEQ ID 8044>. Analysis of this protein sequence reveals the following:

```
Possible site: 46

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3097(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2773

A DNA sequence (GASx1431R) was identified in *S.pyogenes* <SEQ ID 8045> which encodes the amino acid sequence <SEQ ID 8046>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2584(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA98101 GB:M19348 hyaluronidase [Streptococcus pyogenes phage
             H4489A]
 Identities = 337/371 (90%), Positives = 351/371 (93%), Gaps = 1/371 (0%)

Query: 1   MAENIPLRVQFKRMKAAEWASSDVVLLEGKIGFETDTGFAKFGDGQNTFSKLKYLTGPKG 60
            M ENIPLRVQFKRM A EWA SDV+LLEGKIGFETDTGFAKFGDGQNTFSKLKYLTGPKG
Sbjct: 1   MTENIPLRVQFKRMSADEWARSDVILLEGKIGFETDTGFAKFGDGQNTFSKLKYLTGPKG 60

Query: 61  PKGDTGLQGKTGGTGSRGPAGKPGTTDYDQLQNKPDLGAPAQKEETNSKITKLESSKADK 120
            PKGDTGLQGKTGGTG RGPAGKPGTTDYDQLQNKPDLGAPAQKEETNSKITKLESSKADK
Sbjct: 61  PKGDTGLQGKTGGTGPRGPAGKPGTTDYDQLQNKPDLGAPAQKEETNSKITKLESSKADK 120

Query: 121 NAVYLKAESNAKLDEKLNLKGGVMTGQLQFKPN-SGIKPSSSVGGANIDMSKSEGAAMV 179
            +AVY KAES +LD+KL+L GG++TGQLQFKPN SGIKPSSSVGGANIDMSKSEGAAMV
Sbjct: 121 SAVYSKAESKIELDKKLSLTGGIVTGQLQFKPNKSGIKPSSSVGGANIDMSKSEGAAMV 180
```

```
Query: 180 MYTNKDTTDGPLMILRSNKDTFDQSVQFVDYKGTTNAVNIVMRQPTTPNFSSALNITSAN 239
           MYTNKDTTDGPLMILRS+KDTFDQS QFVDY G TNAVNIVMRQP+ PNFSSALNITSAN
Sbjct: 181 MYTNKDTTDGPLMILRSDKDTFDQSAQFVDYSGKTNAVNIVMRQPSAPNFSSALNITSAN 240

Query: 240 EGGSAMQIRGVEKALGTLKITHENPVDKEYDENAAALSIDIVKKQKGKGKTAAQGIYIN 299
           EGGSAMQIRGVEKALGTLKITHENP+V+ +YDENAAALSIDIVKKQKGKGKTAAQGIYIN
Sbjct: 241 EGGSAMQIRGVEKALGTLKITHENPNVEAKYDENAAALSIDIVKKQKGKGKTAAQGIYIN 300

Query: 300 STSGTAGKMLRIRNKNKDKFYVGPDGDPWSCASSIVDGNLTVKDPTSGKHAATKDYVDEK 359
           STSGTAGKMLRIRNKN+DKFYVGPDG F S A+S V GNLTVKDPTSGKHAATKDYVDEK
Sbjct: 301 STSGTAGKMLRIRNKNEDKFYVGPDGGFHSGANSTVAGNLTVKDPTSGKHAATKDYVDEK 360

Query: 360 IAELKKLILKK 370
           IAELKKLILKK
Sbjct: 361 IAELKKLILKK 371
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2774

A DNA sequence (GASx1438R) was identified in *S.pyogenes* <SEQ ID 8047> which encodes the amino acid sequence <SEQ ID 8048>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1892(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

A related DNA sequence <SEQ ID 10439> was identified in GBS which encodes amino acid sequence <SEQ ID 10440>.

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB18711 GB:U38906 ORF36 [Bacteriophage r1t]
 Identities = 70/111 (63%), Positives = 88/111 (79%)

Query: 1   LIEVIIKKYLDEHLDVPSFFEHQKDEPARFIILEKTSGAKQNHLLSSTFAPQSYAESLYE 60
           +IE+IIK +LD HL V SF E + + P  +I+ EKT  +K NHLLSSTFAPQSYA S+YE
Sbjct: 1   MIEIIIKNFLDYHLSVSSFLEKKGEMPLSYILFEKTGSSKSNHLLSSTFAPQSYAPSMYE 60

Query: 61  AALLNDKVKQVIEQLDVLDQVSGVHLNADYNPTDTATKRYRYQAVFDINHY 111
           AA LN+++K+V+E+L  L ++S V LN+DYNPTDT TK YRYQAVFDINHY
Sbjct: 61  AAKLNEQLKEVVERLIELNESISNVSLNSDYNPTDTETKSYRYQAVFDINHY 111
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2775

A DNA sequence (GASx1442R) was identified in *S.pyogenes* <SEQ ID 8049> which encodes the amino acid sequence <SEQ ID 8050>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence
```

```
----- Final Results -----
               bacterial cytoplasm --- Certainty=0.1241(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2776

A DNA sequence (GASx1444R) was identified in *S.pyogenes* <SEQ ID 8051> which encodes the amino acid sequence <SEQ ID 8052>. Analysis of this protein sequence reveals the following:

```
Possible site: 42

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.4547(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2777

A DNA sequence (GASx1447R) was identified in *S.pyogenes* <SEQ ID 8053> which encodes the amino acid sequence <SEQ ID 8054>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                 bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2778

A DNA sequence (GASx1448R) was identified in *S.pyogenes* <SEQ ID 8055> which encodes the amino acid sequence <SEQ ID 8056>. Analysis of this protein sequence reveals the following:

```
Possible site: 20
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3221(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2779

A DNA sequence (GASx1449R) was identified in *S.pyogenes* <SEQ ID 8057> which encodes the amino acid sequence <SEQ ID 8058>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.6356(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2780

A DNA sequence (GASx1453R) was identified in *S.pyogenes* <SEQ ID 8059> which encodes the amino acid sequence <SEQ ID 8060>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2869(Affirmative) < succ>
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2781

A DNA sequence (GASx1455R) was identified in *S.pyogenes* <SEQ ID 8061> which encodes the amino acid sequence <SEQ ID 8062>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1787(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF43512 GB:AF145054 ORF19 [Streptococcus thermophilus
          bacteriophage 7201]
 Identities = 47/126 (37%), Positives = 86/126 (67%), Gaps = 2/126 (1%)

Query: 8   LKDLRNLDLYIASLIRRRDKIEASLL--SSPKWSSDKVNGGIKRKQDVYVELIATAKDI 65
           ++ ++ LD YI S I +  ++E+  L  +S    +D V GG ++ +DD+YVELI   +++
Sbjct: 7   IQQIKALDRYIESQIEQIKRLESQALKVTSGSMHTDMVQGGKRKGKDDIYVELITAREEV 66

Query: 66  EKKTAEAIRKQRSLQNLIDSLENTDSQTILSMVYIDKMTRWQVIDELNCSESTYFRLLRV 125
           E+ TAEAI+++ E +  I ++E+ D++++L MVYID+++ WQ+ D++  S++TY+  LR
Sbjct: 67  ERFTAEAIKQKLESFRRQIANIEDIDARSLLQMVYIDQLSIWQICDFMGISKATYYVKLRQ 126

Query: 126 ATKELN 131
           A K L+
Sbjct: 127 AEKYLD 132
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2782

A DNA sequence (GASx1456R) was identified in *S.pyogenes* <SEQ ID 8063> which encodes the amino acid sequence <SEQ ID 8064>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2883(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB18697 GB:U38906 ORF22 [Bacteriophage r1t]
 Identities = 78/207 (37%), Positives = 123/207 (58%), Gaps = 2/207 (0%)

Query: 6   EIHRILGIDEVYKAPKRLTDILPDKDSREDIPRQFLKYETDVSYDWFMQYPEEEQADRKN 65
           + + +L +DE       R+ ++FDK  RE+ + + L    D+  D+F  YF     A
Sbjct: 7   QFYDMLNVDEHMNFTNRIQELVFDKKGRKEFYSKILNNLHHDMGVDFFRDYFMAHSAVSA- 65

Query: 66  KKQDFTPKSVSTLLSKIISGNQYYEVA-VGTGGILLQAWQEQRLNDSPFTYPSKYWYHV 124
           K Q +TP   +  L + ++ G+   ++   GTG ++IQ WQ+ R+N   F Y PS YWY
Sbjct: 66  KSQHYTPDELSKLTALLVQGSGGALTGAGTGTLIIQKWQDDRMMTDFFNYLPSNYWYQA 125

Query: 125 EELSDKAVPFLLFNMSIRGINGVVVHGDSLTRQVKNIYFLQNTKDDMLSFSDINVMPRTQ 184
            ELSD+A+ FL+   +IRG+NGVV+HGD+L   VK +YF+QN+ ++ + + FS+INV+P ++
Sbjct: 126 LELSDEAISFLIHAFAIRGMNGVVIHGDALEMAVKQVYFIQNSANNPIGFSEINVIPHSK 185
```

```
Query: 185 DIEREPNVKEWIGDGIEHIENPLIEWI 211
           D      + EW    IEHIE+   +WI
Sbjct: 186 DAMEFLGIHEWTRQAIKHIESKPDWI 212
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2783

A DNA sequence (GASx1459R) was identified in *S.pyogenes* <SEQ ID 8065> which encodes the amino acid sequence <SEQ ID 8066>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -2.44   Transmembrane   82 -  98 (  81 -  98)

----- Final Results -----
                bacterial membrane --- Certainty=0.1977(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2784

A DNA sequence (GASx1460R) was identified in *S.pyogenes* <SEQ ID 8067> which encodes the amino acid sequence <SEQ ID 8068>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.3368(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2785

A DNA sequence (GASx1461R) was identified in *S.pyogenes* <SEQ ID 8069> which encodes the amino acid sequence <SEQ ID 8070>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2834(Affirmative) < succ>
```

```
bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2786

A DNA sequence (GASx1462R) was identified in *S.pyogenes* <SEQ ID 8071> which encodes the amino acid sequence <SEQ ID 8072>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3531(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2787

A DNA sequence (GASx1463R) was identified in *S.pyogenes* <SEQ ID 8073> which encodes the amino acid sequence <SEQ ID 8074>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2483(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB14569 GB:Z99117 similar to phage-related protein [Bacillus subtilis]
 Identities = 98/252 (38%), Positives = 152/252 (59%), Gaps = 29/252 (11%)

Query: 16  SPAVKNRIEQVVGARAEQFTTSLLSYISNNNLIAKATSESIMGAAMKAAVINLPIEPSLG 75
           SP+V  R E+V+G RA QFT S+LS+ ++  +L K     S++ +AM AA L+LPI+ +LG
Sbjct: 33  SPSVIKRFEEVLGKRATQFTASILSLYNSEQMLQKTDPMSVISSAMVAATLDLPIDKNLG 92

Query: 76  FAYVVPYNRNYKDGNRWITVNEAQFQIGYRGLIQLAQRSGQVRNIEHGIIYEKEFLGYDK 135
           +A++VPY              +AQFQ+GY+G IQLA R+GQ ++I     I+E E    ++
Sbjct: 93  YAWIVPYG------------GKAQFQLGYKGYIQLALRTGQYKSINCIPIHEGELQKNNP 140

Query: 136 IRGQLKLTGDYVDSGVVKGYPASLRLISGFYKMIPWPKEKVYRHAKKYSKTFDKKTGDFK 195
           +  ++++  +  +S  V GY A  ELI+GF K ++W K +V +H KK+SK+      DF
Sbjct: 141 LTEEIEIDFEKKKSDAVIGYAAYFELINGFRKTVYWTKAQVEKHKKKFSKS------DF- 193
```

```
Query: 196 PGTPWATEPDPMAIKTLLKELLSKYAPLSVEMQDA-LEALNADSTIVIPKDVIPQETNSL 254
           W  ++D MA+KT+LK +LSK+  LSVEMQ A +E D     I    D+T +  +S
Sbjct: 194 ---GWKNDWDAMALKTVLKAVLSKWGILSVEMQKAVIEEDETRERI----DITNEADSS- 245

Query: 255 DDLIGTQNEKKD 266
             ++I ++   KD
Sbjct: 246 -EIIDSEPSNKD 256
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2788

A DNA sequence (GASx1464R) was identified in *S.pyogenes* <SEQ ID 8075> which encodes the amino acid sequence <SEQ ID 8076>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4258(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2789

A DNA sequence (GASx1465R) was identified in *S.pyogenes* <SEQ ID 8077> which encodes the amino acid sequence <SEQ ID 8078>. Analysis of this protein sequence reveals the following:

```
Possible site: 51

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2045(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2790

A DNA sequence (GASx1469R) was identified in *S.pyogenes* <SEQ ID 8079> which encodes the amino acid sequence <SEQ ID 8080>. Analysis of this protein sequence reveals the following:

```
Possible site: 19
```

```
>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2791

A DNA sequence (GASx1470R) was identified in *S.pyogenes* <SEQ ID 8081> which encodes the amino acid sequence <SEQ ID 8082>. Analysis of this protein sequence reveals the following:

```
Possible site: 37

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3577(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC98430 GB:L29324 excisionase [Streptococcus pneumoniae]
 Identities = 23/56 (41%), Positives = 41/56 (73%)

Query: 23 KHLIQQWEGLTVATAKQWATEMRDHPDPKQFVLNPTHRIVFIDYKGPKLFVQWKSR 78
          K ++++W+GL   T  +W  EMR++  F  +V+NPTH++VFI+ +GF+ F++WK +
Sbjct: 19 KGILKRWDGLNKYTLNRWIKEMRENRTFSMYVINPTHKLVFINLEGFESFLRWKQK 74
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2792

A DNA sequence (GASx1473) was identified in *S.pyogenes* <SEQ ID 8083> which encodes the amino acid sequence <SEQ ID 8084>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2725(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2793

A DNA sequence (GASx1476) was identified in *S.pyogenes* <SEQ ID 8085> which encodes the amino acid sequence <SEQ ID 8086>. Analysis of this protein sequence reveals the following:

```
Possible site: 23

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1422(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2794

A DNA sequence (GASx1480R) was identified in *S.pyogenes* <SEQ ID 8087> which encodes the amino acid sequence <SEQ ID 8088>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -4.04   Transmembrane  291 - 307 ( 290 - 309)

----- Final Results -----
             bacterial membrane --- Certainty=0.2614(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2795

A DNA sequence (GASx1489R) was identified in *S.pyogenes* <SEQ ID 8089> which encodes the amino acid sequence <SEQ ID 8090>. Analysis of this protein sequence reveals the following:

```
Possible site: 23

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2278(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2796

A DNA sequence (GASx1490R) was identified in *S.pyogenes* <SEQ ID 8091> which encodes the amino acid sequence <SEQ ID 8092>:

```
SFITSVLAPRKLLKCEGIDLYLMYGDLMTCPEQLLTQLKDWTDVYPNYDESGYGRLRDQKAAQPFKKNGIAVHTYQDHYLHGSQEIINQSG
QPYKVFTPYYRIWQNYPKETPIKVELSQGRWLNLETPDDVLRTVESPKDEKYQDVATFDEASKQLNRPIQDQLAAYHANRDFPAQLGTSRL
SPFLRIGAIGRTVVYHAVRQAPNSLGQATFLKELAWRDFYNMVYVAYPDQKTQPIQKAFSQIEWVNNPDWFQLWKEGKTGYPIVDAAMIQL
QKIGWMHNRLRMIVASPLTKDLLCDWRLGEQYFQQQLIDYDAASNIGGWQWAASTGTDAVFYPRIPNPVTQGKRFDPKGEFIKAYLPQLEH
VPEKYLHEPWKMPKNLQESVSCIIGTDYPQPIVDHAKQREQAIAKYEWAKEKAKIE
```

Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA22361 GB:M94110 DNA photolyase [Bacillus firmus]
 Identities = 175/338 (51%), Positives = 238/338 (66%), Gaps = 6/338 (1%)

Query: 145 EIINQSGQPYKVFTPYYRIWQNYPKETP--IKVELSCGRWLNLETPDDVLRTVES--FKD 200
           +++ + G PYKVFTPYY+ W    K TP  IK ++  G       PD    T+ +    K
Sbjct: 2   QVLKKDGTPYKVFTPYYKAWAKSRKRTPAVIKRDVLLGSVHKGTAPDREAETLFNNLIKK 61

Query: 201 EKYQDVATFDE-ASKQLNRPIQDQLAAYHANRDFPAQLGTSRLSPFLRIGAIGIRTVY-H 258
             Y   A  +E A K+L  F + +L+ Y ANRDFP+  GTSRLSP++ GA+  R++Y H
Sbjct: 62  CSYDWSAIGEEHAIKRLQMPTKKRLSGYKANRDFPSITGTSRLSPYIKTGAVSSRSIYYH 121

Query: 259 AVRQAPNSLGQATFLKELAWRDFYNMVYVAYPDQKTQPIQKAFSQIEWVNNPDWFQLWKE 318
             +    +S    TFLKELAWRDFY MV+   PD K + I + ++ W ++ D    WK
Sbjct: 122 ILNARADSYSAETFLKELAWRDFYRMVHFYEPDCKDREIMEGYRELNWSHDQDDLTSWKR 181

Query: 319 GKTGYPIVDAAMLQLQKTGWMHNRLRMIVASFLTKDLLCDWRLGEQYFQQQLIDYDAASN 378
           G+TG+PIVDA N QL   GWMHNRLRMI ASFLTKDLL DWRLGE+YF++ LIDYD +SN
Sbjct: 182 GETGFPIVDAGMRQLLNEGWMHNRLRMITASFLTKDLLIDWRLGERYFERMLIDYDPSSN 241

Query: 379 IGGWQWAASTGTDAVPYRIFNPVTQGKRFDPKGEFIKAYLPQLEHVPEKYLHEPWKMPK 438
           IGGWQWAAS GTDAVPYRIFNPVTQ KRFD   G +I+ Y+P+L HVP+ Y+HEPWKM +
Sbjct: 242 IGGWQWAASVGTDAVPYRIFNPVTQSKRFDKNGTYIRTYIPELNHVPDHYIHEPWKMSE 301

Query: 439 NLQESVSCIIGTDYPQPIVDHAKQREQAIAKYEWAKEK 476
             Q    C +  DYP PIVDH+KQR++A++ ++   E+
Sbjct: 302 EEQVKYKCRLDEDYPLPIVDHSKQRKKALSFPKGDDEE 339
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2797

A DNA sequence (GASx1493R) was identified in *S.pyogenes* <SEQ ID 8093> which encodes the amino acid sequence <SEQ ID 8094>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2748(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2798

A DNA sequence (GASx1501R) was identified in *S.pyogenes* <SEQ ID 8095> which encodes the amino acid sequence <SEQ ID 8096>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -7.27   Transmembrane   64 -  80 (  53 -  83)

----- Final Results -----
                bacterial membrane --- Certainty=0.3909(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC95443 GB:AF068901 YlmG [Streptococcus pneumoniae]
 Identities = 35/81 (43%), Positives = 58/81 (71%)

Query: 1  MILILSILLRLIKVYTYLLIAYALMSWFPGAYDSKIGRLISGIVEPILKPFRAFNLQFAG 60
          MI ++ ++    + +Y+ +L+A+A+MSWFPGAY+S +GR I  +V+P+L P +   LQ AG
Sbjct: 1  MIFLIRMIYNAVDIYSLILVAFAVMSWFPGAYSSSLGRWIVALVKPVLAPLQRLPLQIAG 60

Query: 61 LDFTIFVVIISLNFLAQVLVR 81
          LD +++V I+ + FL + LVR
Sbjct: 61 LDLSVWVAIVLVRFLGENLVR 81
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2799

A DNA sequence (GASx1502) was identified in *S.pyogenes* <SEQ ID 8097> which encodes the amino acid sequence <SEQ ID 8098>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -2.39   Transmembrane   17 -  33 (  17 -  33)

----- Final Results -----
                bacterial membrane --- Certainty=0.1956(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2800

A DNA sequence (GASx1507) was identified in *S.pyogenes* <SEQ ID 8099> which encodes the amino acid sequence <SEQ ID 8100>. Analysis of this protein sequence reveals the following:

```
Possible site: 23

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.0865(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2801

A DNA sequence (GASx1511R) was identified in *S.pyogenes* <SEQ ID 8101> which encodes the amino acid sequence <SEQ ID 8102>. Analysis of this protein sequence reveals the following:

```
Possible site: 47

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-11.83   Transmembrane    31 -  47 (  22 -  53)
   INTEGRAL    Likelihood = -0.96   Transmembrane     2 -  18 (   1 -  18)

----- Final Results -----
              bacterial membrane --- Certainty=0.5734(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2802

A DNA sequence (GASx1516R) was identified in *S.pyogenes* <SEQ ID 8103> which encodes the amino acid sequence <SEQ ID 8104>. Analysis of this protein sequence reveals the following:

```
Possible site: 42

>>> Seems to have no N-terminal signal sequence
```

```
----- Final Results -----
                bacterial cytoplasm --- Certainty=0.2729(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:HAA96472 GB:AB036428 Dpr [Streptococcus mutans]
 Identities = 132/175 (75%), Positives = 153/175 (87%)

Query: 1   MTNTLVENIYASVTHNISKKEASKNEKTKAVLNQAVADLSVAASIVHQVHWYMRGPGFLY 60
            MTNT+ ENIYAS+ H + KKE S NEKTKAVLNQAVADLS AASIVHQVHWYMRG GFLY
Sbjct: 1   MTNITTENIYASIIHQVEKKENSGNEKTKAVLNQAVADLSKAASIVHQVHWYMRGSGFLY 60

Query: 61  LHPKMDELLDSLNANLDEMSERLITIGGAPYSTLAEFSKHSKLDEAKGTYDKTVAQHLAR 120
            LHPKMDEL+D+LN +LDE+SERLITIGGAP+STL EF ++S+L+E  GT+DK++  HL R
Sbjct: 61  LHPKMDELMDALNGHLDEISERLITIGGAPFSTLKEFDENSKLEETVGTWDKSITDHLKR 120

Query: 121 LVEVYLYLSSLYQVGLDITDEEGDAGTNDLFTAAKTEAEKTIWMLQAERGQGPAL 175
            LV+VY YLSSLYQVGLD+TDEE DA +ND+FTAA+TEA+KTIWMLQAE GQ P L
Sbjct: 121 LVQVYDYLSSLYQVGLDVTDEEDDAVSNDIFTAAQTEAQKTIWMLQAELGQAPGL 175
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2803

A DNA sequence (GASx1517) was identified in *S.pyogenes* <SEQ ID 8105> which encodes the amino acid sequence <SEQ ID 8106>. Analysis of this protein sequence reveals the following:

```
Possible site: 46

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -6.32   Transmembrane  109 - 125 ( 106 - 126)
   INTEGRAL    Likelihood = -5.26   Transmembrane   63 -  79 (  61 -  81)
   INTEGRAL    Likelihood = -5.20   Transmembrane  154 - 170 ( 151 - 176)
   INTEGRAL    Likelihood = -4.14   Transmembrane  189 - 205 ( 189 - 205)
   INTEGRAL    Likelihood = -3.50   Transmembrane  130 - 146 ( 127 - 147)
   INTEGRAL    Likelihood = -2.92   Transmembrane    6 -  22 (   1 -  24)
   INTEGRAL    Likelihood = -2.23   Transmembrane   83 -  99 (  83 - 101)

----- Final Results -----
                bacterial membrane --- Certainty=0.3527(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:HAA96471 GB:AB036428 type IV prepilin peptidase homologue
             [Streptococcus mutans]
 Identities = 55/127 (43%), Positives = 78/127 (61%), Gaps = 3/127 (2%)

Query: 83  VSASYCYLLLFSLLPSLFDWRSQEYPPILWLFSFVSLLLFYSINYLSLILLLIGLLAHLR 142
            ++ S   LL   +L SL+D + Q YP  LW+      L+  Y +N +SLIL L G+ A L+
Sbjct: 91  LITSQVCLLFMGVLLSLYDLQDQSYPLTLWIGFTFLLMFIYPLNLISLIILFLFGIFAALK 150

Query: 143 PFSIGAGDFFYLASLALVLDLTSLIWLIQLASLAGITACLLLGIKRIP--FIPYLSPGLF 200
              +IG+GDFFYLA+LAL L+L  +IW+IQ+ASL GI   LL     + P  P+P+L  G
Sbjct: 151 NINIGSGDFFYLATLALSLNLQQIIWIIQIASLLGILYSLLFQKHKEPFAPVPFLPLG-H 209

Query: 201 WIVLLEH 207
```

```
          I++  H
Sbjct: 210 LIIIFSH 216
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2804

A DNA sequence (GASx1538R) was identified in *S.pyogenes* <SEQ ID 8107> which encodes the amino acid sequence <SEQ ID 8108>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1186(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2805

A DNA sequence (GASx1539R) was identified in *S.pyogenes* <SEQ ID 8109> which encodes the amino acid sequence <SEQ ID 8110>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have an uncleavable N-term signal seq
   INTRGRAL    Likelihood =-11.73   Transmembrane    6 -  22 (   3 -  32)

----- Final Results -----
              bacterial membrane --- Certainty=0.5692(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF31453 GB:AF221126 putative histidine kinase [Streptococcus pneumoniae]
 Identities = 141/301 (46%), Positives = 210/301 (68%), Gaps = 7/301 (2%)

Query: 1   MKRYPLLVQLISYVFVIVIALITTLGLLVYQTSSRNIRQLIRRDTRQSIRQSSQFIDAYI 60
            MKR  LLV+++   +P++ + L+  +G  YYQ+SS  I   IE +++ +I Q+S FI +YI
Sbjct: 1   MKRSSLLVRMVISIFLVFLILLALVGTFYYQSSSSAIEAIIEGNSQTYISQTSHFIQSYI 60

Query: 61  KPLKETTSVLAKNTEIQAFASQIHQENDKQVLQLMKMVLATNSDLQAAVLVTKDGRTVST 120
            K L+ T++ L + T++ A+A    Q+  + +  L   +L ++ DL+  VLVTK G+ +ST
Sbjct: 61  KKLETTSTGLTQQTDVLAYAENPSQDKVEGIRDLFLTILKSDKDLKTVVLVTKSGQVIST 120

Query: 121 NSQQITMKTSSDMMAEEWYKAAIDRQAMPILTPARQLSLSSKKEWVVSVTQRVVDRAGHNL 180
            +  + MKTSSDMMAE WY+ AI + AMP+LTPAR+     S  +WV+SVTQR+VD  G NL
Sbjct: 121 DDSVQMKTSSDMMAEDWYQKAIHQGAMPVLTPARK----SDSQWVISVTQRLVDAKGANL 176

Query: 181 GVLRLDIAYPTIKASLDQLQLGRQGFAFIVNDKHEFVYHPKKSVYSSSKEMAAMKPYLAI 240
            GVLRLDI+Y T++A L+QLQLG+QGFAFI+N+ HEFVYHP+ +VYSSS +M AMKPY+
```

```
Sbjct: 177 GVLKLDISYETLSAYLNQLQLGQQGFAFIINENHEFVYHPQHTVYSSSSKMEAMKPYIDT 236

Query: 241 QNGYTKDKTSFVVQKLIPNSQWTLVGVASLDQLHRVQRQIFWSFSWNRASTLSDLWLCNCL 301
           GYT    S+V Q+ I  + WT++GV+SL++L +V+ Q+ W+      ++++ L +C CL
Sbjct: 237 GQGYTPGHKSYVSQSKIAGTDWTVLGVSSLEKLDQVRSQLLWTL---LGASVTSLLVCLCL 294
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2806

A DNA sequence (GASx1542R) was identified in *S.pyogenes* <SEQ ID 8111> which encodes the amino acid sequence <SEQ ID 8112>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> May be a lipoprotein

----- Final Results -----
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC23101 GB:U32823 conserved hypothetical protein [Haemophilus influenzae Rd]
 Identities = 56/128 (43%), Positives = 87/128 (67%)

Query: 73  DFELKGIDGKTYRLSEFKGKKVYLKFWASWCSICLSTLADTEDLAKMSDKDYVVLTVVSP 132
           D +LK ++ +   LS++KGK VY+K WASWC ICL+ LA+ +DL+   D+++ V+T+VSP
Sbjct: 24  DVQLKDLNNQPVTLSQYKGKPVYVKWWASWCPICLAGLAEIDDLSAEKDRNFEVITIVSP 83

Query: 133 GHQGEKSEADFKKWFQQTDYKDLPVLLDPDGKLLEAYGVRSYPTTEVFIGSDGVLAKGHIG 192
            H+GEK  ADF +W++G +YK++ VLLD  G+++    VR YP  +F+ SD  L K   G
Sbjct: 84  DHNGEKDTADFTEWYKGLEYKNITVLLDEKGEIIDKARVRGYPFNLFLDSDLNLKKTVFG 143

Query: 193 YAKKSDIK 200
           +     I+
Sbjct: 144 HLGAEQIR 151
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2807

A DNA sequence (GASx1543R) was identified in *S.pyogenes* <SEQ ID 8113> which encodes the amino acid sequence <SEQ ID 8114>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL     Likelihood = -7.75   Transmembrane  171 - 187 ( 169 - 191)
   INTEGRAL     Likelihood = -6.26   Transmembrane  205 - 221 ( 203 - 232)
   INTEGRAL     Likelihood = -5.73   Transmembrane   56 -  72 (  54 -  81)
   INTEGRAL     Likelihood = -5.36   Transmembrane   92 - 108 (  91 - 113)
   INTEGRAL     Likelihood = -3.45   Transmembrane   20 -  36 (  14 -  39)
   INTEGRAL     Likelihood = -1.17   Transmembrane  147 - 163 ( 144 - 163)

----- Final Results -----
                bacterial membrane --- Certainty=0.4100(Affirmative) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC23102 GB:U32823 cytochrome C-type biogenesis protein
            [Haemophilus influenzae Rd]
 Identities = 106/224 (47%), Positives = 138/224 (61%), Gaps = 16/224 (7%)

Query: 6   VLMVSVPGAGLLSFPSPCIFPVLPVYILGILLDADDSKTITPGKKLYWYGIVKTLAFIPG 65
           +L+ +VF AGL SF SPCIFP++P+Y GIL              GKK     ++ T  FI G
Sbjct: 6   LLIGTVFLAGLASFLSPCIFPILPIYFGILSKG---------GKK-----VLMTFLFILG 51

Query: 66  LSTIPVILGYGAGFLGNILVAVWFRYLLGALVIILGIHQMGLITIKSLQFQKSLTPHNNK 125
           LS  FV LG+  GFLGNIL++    R + G +VIILGIHQ+G+  I  L+  K +     +
Sbjct: 52  LSLTFVSLGFSFGFLGNILPSNTTRILAGVIVIILGIHQLGIFKIGLLERTKLVEIKTSG 111

Query: 126 NRNGLFNAFILGLTFSFGWTPCVGPVLSSVIALVASGGNGAWQKSVLMIIYTLGIGIPFL 185
                L  AF+LGLTFS GWTPC+GP+L+SVIAL    G+ A  G  +M +Y LGL  PF+
Sbjct: 112 KSTAL-EAFVLGLTFSLGWTPCIGPILASVIALSGDEGS-ALYGASMMFVYVLGLATPFV 169

Query: 186 LISFASGIVLKQFNKLKPHILLLKKVGGVLIIVMGILLMTGTLN 229
           L SF S  +LK+   L  H+   K  GG+LIIVMGILL+T   +
Sbjct: 170 LPSFFSDSLLKRAKGLNKHLDKFKIGGGILIIVMGILLITNNFS 213
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2808

A DNA sequence (GASx1544) was identified in *S.pyogenes* <SEQ ID 8115> which encodes the amino acid sequence <SEQ ID 8116>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1493(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2809

A DNA sequence (GASx1546R) was identified in *S.pyogenes* <SEQ ID 8117> which encodes the amino acid sequence <SEQ ID 8118>. Analysis of this protein sequence reveals the following:

```
Possible site: 46

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4658(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04061 GB:AP001508 unknown conserved protein in others
          [Bacillus halodurans]
 Identities = 48/89 (53%), Positives = 61/89 (67%)

Query: 1  MMVLVTYDVNTETPAGRKRLRHVAKLCVDYGQRVQNSVFECSVTPARFVDIKHRLTQIID 60
          M+VL+TYDV T +  G KRLR VAK C +YGQRVQNSVFEC V    +   +K  LT +ID
Sbjct: 1  MLVLITYDVQTSSMGGTKHLRKVAKACQNYGQRVQNSVFECIVDSTQLTSLKLELTSLID 60

Query: 61 EKTDSIRFYLLGKNWQRRVETLGRSDSYD 89
          E+ DS+R Y LG N++ +VE +G   S D
Sbjct: 61 EEKDSLRIYRLGNNYKTKVEHIGAKPSID 89
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2810

A DNA sequence (GASx1547R) was identified in *S.pyogenes* <SEQ ID 8119> which encodes the amino acid sequence <SEQ ID 8120>. Analysis of this protein sequence reveals the following:

```
Possible site: 57

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.70   Transmembrane   44 -  60 (  43 -  60)

----- Final Results -----
               bacterial membrane --- Certainty=0.1680(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>

RGD motif: 330-332
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04060 GB:AP001508 unknown conserved protein in others
          [Bacillus halodurans]
 Identities = 162/341 (47%), Positives = 231/341 (67%), Gaps = 1/341 (0%)

Query: 1   MKKLLNTLYLTQEDFVVTKEGDNIVIKQEGKVLKRPPFRIIDGIVCFSYLGVSSALVKLC 60
           MKKLLNTLY+TQ D Y++ +GDN+V+ +E + L R P   ++ IV F Y G S AL+  C
Sbjct: 1   MKKLLNTLYVTQEDTYLSLDGDNVVLLKEQEKLGRLPLHNLEAIVGFGYT>FEATURESALMGYC 60

Query: 61  TENQINLSPHTPQGRFCGRYIGSTNGNVLLRRHHYRLSDRE-RSLEYAKRFILAKISNSR 119
            E  I+++F T  GRF  R +G + GNV+LR+  YR+S+ + RS + A+ FI  K+ NS+
Sbjct: 61  AERNISITFLTKNGRFLARVVGESRGNVVLRKTQYRISENDQESTYIARNFITGKVYNSK 120

Query: 120 KYLLRFKRDHRQQIDTKLPEAVNDELIWALSMVQAADNKDSLRGISGQAANQYFRIFNDL 179
             L R  R+H  +++ + F+A +  L   ++ ++  D+ +SLRG SGQAA  Y ++F+ +
Sbjct: 121 WMLERMTRESHLRVNVEQFKATSQLLSVMMQEINNCDSLESLRGWSGQAAINYNKVFDQM 180

Query: 180 VLTDKKTFYFQGRSKRPPLDCVNALLSFGYSLLTFECQSALEAVGLDSYVGFPHTDRPGR 239
           +L  K+ F F GRS+RPP D VNA+LSF Y+LL  +ALE VGLD+YVGF H DRPGR
Sbjct: 181 ILQQKEEFAFHGRSRRPPKDNVNAMLSFAYTLLANDVAAALETVGLDAYVGFMHQDRPGR 240

Query: 240 ASLALDLVEEFRSYIVDRFVFSLINKGQLQKGHFEVKENGSILLTSNGRAIFIDLMQKRK 299
           ASLALDL+EE R    DRFV SLIN+ ++    F  KENG++L+T+  R  F+  WQ +K
Sbjct: 241 ASLALDLMEELRGLYADRFVLSLINRKEMTADGFVKKENGAVLMTDEARKTFLKAWQTKK 300

Query: 300 HTEVEHPFTKEKVKLMLLPVVQAQLLAKAIRGDLESYPPFM 340
           ++ HP+  EK+   L+PVVQA LLA+ +RGDL+ YPPF+
Sbjct: 301 QEKITHPYLSEKMSWGLVPYVQALLLARPLRGDLDEYPPFL 341
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2811

A DNA sequence (GASx1548R) was identified in *S.pyogenes* <SEQ ID 8121> which encodes the amino acid sequence <SEQ ID 8122>. Analysis of this protein sequence reveals the following:

```
Possible site: 49

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2247(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04059 GB:AP001508 unknown [Bacillus halodurans]
 Identities = 90/169 (53%), Positives = 111/169 (65%), Gaps = 1/169 (0%)

Query: 45  LHTKADNPYIKEKRKELLVSRAMPISSARIGLSGIMDVVEPYKDDQGVSLRGKRGKWLPK 104
           +H KAD P++KEKR   L  RAMPI S  L +SGI DVVEP +D +G+ L G  G +
Sbjct: 1   MHKKADQPFMKEKRGSKLTVRAMPIQSKNLQISGICDVVEPVQDSEGIELSGVSGSYKAF 60

Query: 105 VVEYKRGKPKKDTRDIVQLVAQTWCLEETLDCDINSGCLYYISVNQRVIVPMTSALRQEV 164
            VEYKRGKPKK   DIVQLVAQ MCLEE L C I++G L+Y+ +  RV VP+T ALR +V
Sbjct: 61  PVEYKRGKPKKGDEDIVQLVAQAMCLEEMLVCRIDKGYLFYNEIKHRVEVPITDALRDKV 120

Query: 165 KRLAAEMHEVVQSQMLPKAAYFKNCQLCSLVDICKPRLSKKTESVSKYI 213
            ++A EMH  Y+++  PK      C  CSL  IC P+L  K RSV KYI
Sbjct: 121 VQMAKEMHHYYSNKHTPKVKTGPPCNNCSLQSICLPKLMNK-RSVKKYI 168
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2812

A DNA sequence (GASx1549R) was identified in *S.pyogenes* <SEQ ID 8123> which encodes the amino acid sequence <SEQ ID 8124>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1399(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04058 GB:AP001508 unknown conserved protein in others
            [Bacillus halodurans]
 Identities = 148/290 (51%), Positives = 190/290 (65%), Gaps = 19/290 (6%)

Query: 6   MLEHKIDFMVTLEVKEANANGDPLNGNMPRTDAKGYGVMSDVSIKRKIRNRLQDMGKSIF 65
```

```
            +L+HKIDF V L V +AN NGDPLNGN PR +  G+G +SDV+IKRKIRNRL DM + IF
Sbjct: 3    LLDHKIDFAVILSVTKANPNGDPLNGNRPRQNYDGHGEISDVAIKRKIRNRLLDMEEPIF 62

Query: 66  VQANERIEDDFRSLEKRFSQH-----FTAQTPDKEIEEKANAL---WFDVRAKQCVFTYLK 118
            VQ+++R  D P+SL  R   +       K  +  ++E A      W DVR+FGQVF +
Sbjct: 63  VQSDDRKADSPKSLRDRADSNPELAKMLKYANASVDEFAKIACQEWMDVRSFGQVFAFKG 122

Query: 119 K--SIGVRGPVSISMAKSLEPIVISSLQITRSTNGMEAKNNSGRSSDTMGTKHFVDYGVY 176
               S+GVRGPVSI  A S++PI I S QIT+S N +       RSSDTMG KH VD+GVY
Sbjct: 123 SNLSVGVRGPVSIHTATSIDPIDIVSTQITKSVNSVTGDK---RSSDTMGKHVDFGVY 179

Query: 177 VLKGSINAYFAEKTGFSQRDAEAIKEVLVSLFENDASSARPBGSMRVCEVFWPTHSSKLG 236
            V KGSIN   AEKTGF+ EDAE IK  L++LFENDS+SSARP+GSM V +V+W+ HSSKLG
Sbjct: 180 VFKGSINTQLAEKTGFTNEDAEKIKRALITLFENDSSSARPDGSMEVHKVYWWEHSSKLG 239

Query: 237 NVSSARVFDLLEYHQSIEEKSTYDAYQIHLNQEKLAKYEAKGLTLELILEG 286
            SSA+V   L+    +  ++D Y + L       YE  GL +E++++G
Sbjct: 240 QYSSAKVHRSLKIESKTDTPKSFDDYAVEL-------YSLDGLSVEVIDG 282
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2813

A DNA sequence (GASx1550R) was identified in *S.pyogenes* <SEQ ID 8125> which encodes the amino acid sequence <SEQ ID 8126>. Analysis of this protein sequence reveals the following:

```
Possible site: 43

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2882(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04057 GB:AP001508 unknown [Bacillus halodurans]
 Identities = 176/671 (26%), Positives = 311/671 (46%), Gaps = 87/671 (12%)

Query: 1   MDFFTSLLKTYEKAELADLVDHQKR--NNEPVLLPIYHTSLKSNGKNIISVKLDKDGQFH 58
            M +   L +TYE A L  +    K+  + E  LLPI HT+  ++    I V LD+DG F
Sbjct: 1   MSWLLHLYETYE-ANLDQVGKTVKKGEDRSYTLLPISHTIQNAH----IEVTLDEDGDFL 55

Query: 59  KAEPMADKQMIIFPVTADSVARSGSHPAPHPLVDKPAYYSAEN----GQIQ------YDS 108
            +A+ +  K+  + P T ++ +RSGS  AP+PL DK +Y + +     G+I+      +D+
Sbjct: 56  RAKALT-KESTLIPCTEEAASRSGSKVAPYPLHDKLSYVAGDPVKYGGKIKNQDDAPFDT 114

Query: 109 FHKQNNWID--YCERGDVKKFLTPVQQFILKPEFLTLILDSLIGPDYQHNQLKVTFCDA 166
            + K L  W +  Y  E  VK   T++++  L  + +  L       NQ  +   +
Sbjct: 115 YIKNLGEWANSPYATE-EVKCIYTYLKKGRLIEDLVDAGVLKL-----DENQQLIKWEEK 168

Query: 167 TGKEKLIDLSACFLEPSIDQ------FQGFKNESVSTF---KALHQSYISFVRANRENLG 217
             +E L +  A F   + DQ      F  F  ES+       K +  S+ISF
Sbjct: 169 RYEELLGEKPAIFSSGATDQASAFVRFNVFHPESIDDVWRDKEMFDSFISFYNDKLGEED 228

Query: 218 ICNISGRKEQLQTKH----RGLMQNAKIISVS-NKREAYKGRFERRDVFSVGYETSEKI 272
            IC ++G    T++H    R     AK+IS + N     ++GRF+    +  + YE S+K
Sbjct: 229 ICFVTGNRLPSTERHANKIRHAADKAKLISANDNSGFTFRGRFKTSREAVGISYEVSQKA 288

Query: 273 HLMLKYLLENKNTSTWLGSSQYLINWPSDD-LTNDSRLDVSPIFDDGLEDDDDDTPPV 331
            H  LK+L+  ++ S     +  + W +D+ L  +  D V  +    E + D DT  +
Sbjct: 289 HNALKWLIHRQSKSI---DDRVPLVWSNDNSLVPNPDEDAVDIMKHANRELERDPDTGQI 345
```

```
Query: 332 ITLATKDNKRIGKSFIKGQKLFANDATY----YVAILNKTSNGRIALKYFRQLQASQLLT 387
           A E  K IG         + +D  Y    ++ +L+  + GR+A+ Y+R L      L
Sbjct: 346 F--AGEVKKAIGG--------YRSDLNYQPHVHILVLDSATGRMAVLYYRSLNKELYLN 395

Query: 388 NLNKWQETTYSMESRSKFGKSRLRT-----PTFHDLLNVSYGVDRDRFLELDNDNFKSDQIQ 443
            L  W ++ +WE R +  +    +    P   DI   +YG            ++    D ++
Sbjct: 396 RLEAWHDSCAWESHRYRRDEKEFISFYGAPATKDIAFAAYGPRA-------SEKVIKDLME 448

Query: 444 KLVASLIDGKDMPQSIVKKL---GNNVKERHRYRKHWVQVEQVCLAILHK---QNGEEFS 497
           +++  ++DG+ +P+ IV+       +N     R+   W +   + A++ K   + EE+
Sbjct: 449 RMLPCIVDGRRVPKDIVRSAFQRASNPVSMRRWK--WKKTLSITCALIRKMHIEQKEEWG 506

Query: 498 PMLDHTNQNRSYLFGRLLAIFELIETLRYGLDGNNNDRITNAERYWTAYTGQPTKLMMLL 557
            LD ++ +RSYLFGRLLA+ +++E    G  G +  R TNA RY  +Y+  P +     +
Sbjct: 507 VPLDKSSTDRSYLFGRLLAVADVLER---GALGKDETRATNAIRYMNSYSKNPGRTWKTI 563

Query: 558 ENKLKPYEEPLKLANRGSWMKLEKKKEILELLNPLLETETMEKPLDVRFLRGYYAEKNY 617
           +  ++FY+   KL  + ++  L K  +EI +    P   +    PL  +++ G+Y+++ +
Sbjct: 564 QESLQFYQ--AKLGTKATY--LGKLVDEIGDQFEP---GDPNNMPLTSQYLLGFYSQRRE 616

Query: 618 YYTKQNTEVTE 628
            Y K+  E  +
Sbjct: 617 LYKKKKEETNQ 627
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2814

A DNA sequence (GASx1551R) was identified in *S.pyogenes* <SEQ ID 8127> which encodes the amino acid sequence <SEQ ID 8128>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3035(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04056 GB:AP001508 unknown [Bacillus halodurans]
 Identities = 90/218 (41%), Positives = 127/218 (57%), Gaps = 7/218 (3%)

Query: 13  GQRALFTNPATKGGSERSSYSVPTRQALNGIVDAIYYKFPTTNIVTEVKVINQIQTELQG 72
           G  ALFT+P TK G E+ SYSVPT QAL GI ++IY+KPT   ++ E++V+  IQ E +G
Sbjct: 11  GDYALFTDPLTKIGGEKLSYSVPTYQALKGIABSIYWKPTIVFVIDELRVMKPIQMESKG 70

Query: 73  VRALLHDYSADLSYVSYLSDVVYLIKPHPVWNEDRKDLNSDRLPAKHRAIMERSIRKGGR 132
           VR + +     L++ +YL DV Y +K HF +N  R DL  DR   KH +I++RS++ GGR
Sbjct: 71  VRPIEYGGGNTLAHYTYLKDVHYQVKAHFEFNLHRPDLAFDRNEGKHYSTLQRSLKAGGR 130

Query: 133 RDVFLGTRECLGLVDDISQESVETTVSYYNGV-NIDLGIMFHSFAYPKDK-KTPLKSYFT 190
            RD+FLG RSC G V   + E+ +   +Y+G     LG M H F YP +  +  L
Sbjct: 131 RDIFLGSRECQGYV---APCSPGSGDGFTDSQGKVHLGTMVHGFNYPDETQQHQLDVRLW 187

Query: 191 KTVMKNGVITFKAQSECDIVNTLSSYAFKA--PEEIKS 226
            VM+NG I F    +C IV  +     K   P+ ++S
Sbjct: 188 SAVMENGYIQFPRPEDCPIVRPVKEMEPKIFNPDNVQS 225
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2815

A DNA sequence (GASx1552R) was identified in *S.pyogenes* <SEQ ID 8129> which encodes the amino acid sequence <SEQ ID 8130>. Analysis of this protein sequence reveals the following:

```
Possible site: 53

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2770(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04055 GB:AP001508 unknown conserved protein in others
            [Bacillus halodurans]
 Identities = 252/836 (30%), Positives = 404/836 (48%), Gaps = 90/836 (10%)

Query: 3   MILAHYDCKKDKKQSLDEHLWHVACSSRQEASIIGQGDVLPLIGLYHDLGKADRTFQD-- 60
           M +AH         Q+L EHL  V C +     + +     V  L GL HDLGK     F+D
Sbjct: 1   MYIAHIREVDKVIQILKEHLCGVQCLAETPGAKLRLQHVAGLAGLLHDLGKYTNEFKDYI 60

Query: 61  -------KLLNNPNRHVDHSYAGAKYLCSIIGPHLKNRGVDKNERMTFNEMVGYVISAHH 113
                 +L          VDHS AG + L  +     L +R       +E++   E+VG   I +HH
Sbjct: 61  YKAVFEPELAEKKRGQVDHSTAGGRLLYQM----LHDRENSPHEKLL-LAEVVGNAIISHH 115

Query: 114 GMYDLCYYFDDAEVYGPNKFKNRINRDLDGYHYHEDIKGYALKLEKKLCDYGYK-DLREL 172
                    +Y       N + R L+      +++  Y   +E+    +  + +L
Sbjct: 116 SNLQ--------DYISPTIESNFLTRVLE-----KELPEYESAVERFFQEVMTEAELARY 162

Query: 173 IDKAFDNYQQAMSSLNWQDKSEWDYYQSCMVRLVLSLLKNADILDTVNAYGLKISPMDKT 232
                + KA D  +Q   +   Q       Y  SC++         +AD  +T   +  +    + T
Sbjct: 163 VAKAVDEIKQPTDNSPTQSFFLTKYIFSCLI--------DADRTNT-RMFDEQAREEEPT 213

Query: 233 ERSFLKHSYLAAIEQKYASPSGQPNNQ---LNTIRTEIAERVKERGKRDSKGIYRLDLPTG 289
                +   L   Y   +   AS  + ++     +N +R+ ++E+ +     R S GIY L +PTG
Sbjct: 214 QPQQLFEHYHQQLLNHLASLKSSDSAQKPINVLRSAMSEQCESFAMRPS-GIYTLSIPTG 272

Query: 290 AGKTNLSMRYAFHQLVHHDKSRFFVITPFLSVLEQNASEIRKVTGD-LGVLEHHSNVVKQ 348
             GKT  S+RYA      ++K R  YI PF +++EQNA E+R + GD   +LEHHSNVV+
Sbjct: 273 GGKTLASLRYALKHAQEYNKQRIIYIVPFTTIIEQNAQEVRNILGDDENILEHHSNVVED 332

Query: 349 ANEDDDDKDSLLSA-----YLSDSWDSQVVLTSMVQFFQTLFKTKSANLRRFSSLINSVV 403
                +   D+ +D +++          D+WD  ++ T++VQF     + N RR  +L +SV+
Sbjct: 333 SENGDEQEDGVITKKERLRLARDNWDRPIFTTLVQFLNVFYAKGNRNTRRLHNLSHSVL 392

Query: 404 ILDEVQSLPIEVTTLFNLTMNFLNKVMDTTTVLCTATQPAYDSSEIDHRICYGGNLGSLA 463
                I DEVQ +P +  +LFN  +NFL +     +I+LCTATQP  ++  + H +     +
Sbjct: 393 IFDEVQKVPTKCVSLFNEALNFLKEFAHCSILLCTATQPTLEN--VKHSLLKDRD----G 446

Query: 464 EIVELTIEEKQIPSRTELRKFDDSDQKVLTDVINLILGEE---NSVLAIFNTKKTVINC 520
                EIV+   E  + F R E+   D +DQ +     +  +  E    S L I NTKK V +
Sbjct: 447 EIVQNLTSVSEAFKRVEI--LDKTDQPWTNERLAEWVRDEAPSWGSTLIILNTKKVVKDL 504

Query: 521 YTMLKDMTDRPVYQLSTNMCAQHRLDLIAKIKTELQNNIPIICISTQLIKAGVDVDFHRV 580
                Y  L+      PV+ LST+MCA HR D + +I+  L+    P IC++TQLIKAGVDV F  V
Sbjct: 505 YEKLEG-GPLPVFHLSTSMCAAHRKDQLDEIRALLKRGTPFICVTTQLIKAGVDVSFKCV 563

Query: 581 IRSYSGIDSIVQAAGRCNREGKRDKGQVTLVNLTNEERNISRLTEIKTKKEATSSILHKI 640
```

```
            IRS +G+DSI QAAGRQNR G+     V +++  + EE +S+L EI+  +E   ++L +
Sbjct: 564 IRSLAGLDSIAQAAGRQNRHGKEQLQYVYVID--HASETLSKLKEIRVGQEIAGNVLARF 621

Query: 641 GSPIDISTLN-------RDFPEYYYANNQGLMDYPLED-----NLSIYDYLSLNIYQTAN 688
               +    N        R++F YYY+      ++Y +++          +   +   N Y T
Sbjct: 622 KKKASKYEGNLLSQAAMKEYFKYYYSKMDANLNYFVKEVDKDMTKLLMSHAVENSYVTYY 681

Query: 689 KKFKGK-----LKQAFKTAGAKMNLINNDMIGILVPYGEAKKLAYLKELGVSHPLSAKD 743
           +K  G      L  ++KTA      +I+ +     +VPYGE +  +A L          S +
Sbjct: 682 QKNIGTEFPLLLNGSYKTAADHFRVIDQNTTSAIVPYGEQQDIAQLN--------SGEW 733

Query: 744 YQTIKSLLKELQPFTVNV--RENDPLFE--TFKSYLNGQILVLTSEYYDTERGVKY 795
               +  +LK+ Q +TVN+  +E D L +      +L+G +  L   +Y  + GV +
Sbjct: 734 VDDLSKVLKKAQQYTVNLYSQEIDQLKKEGAIVMHLDGMVYELKESWYSHQVGVDF 789
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2816

A DNA sequence (GASx1558) was identified in *S.pyogenes* <SEQ ID 8131> which encodes the amino acid sequence <SEQ ID 8132>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1050(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2817

A DNA sequence (GASx1563) was identified in *S.pyogenes* <SEQ ID 8133> which encodes the amino acid sequence <SEQ ID 8134>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1872(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2818

A DNA sequence (GASx1564R) was identified in *S.pyogenes* <SEQ ID 8135> which encodes the amino acid sequence <SEQ ID 8136>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2173(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2819

A DNA sequence (GASx1566R) was identified in *S.pyogenes* <SEQ ID 8137> which encodes the amino acid sequence <SEQ ID 8138>. Analysis of this protein sequence reveals the following:

```
Possible site: 43

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3486(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2820

A DNA sequence (GASx1568) was identified in *S.pyogenes* <SEQ ID 8139> which encodes the amino acid sequence <SEQ ID 8140>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2711(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2821

A DNA sequence (GASx1569) was identified in *S.pyogenes* <SEQ ID 8141> which encodes the amino acid sequence <SEQ ID 8142>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2822

A DNA sequence (GASx1576R) was identified in *S.pyogenes* <SEQ ID 8143> which encodes the amino acid sequence <SEQ ID 8144>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4042(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2823

A DNA sequence (GASx1577R) was identified in *S.pyogenes* <SEQ ID 8145> which encodes the amino acid sequence <SEQ ID 8146>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3342(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04515 GB:AP001509 unknown [Bacillus halodurans]
 Identities = 36/104 (34%), Positives = 55/104 (52%)
```

```
Query: 2    HMGAWNTGNNKILYTQESVTDDMIAKRDQSIKDAKESPILGFTVDTKVIKTKLSNISNVM 61
            +M ++  GN  IL   E    D     +   + A  SP LGF  D+  ++TE++ ISNV
Sbjct: 392 NMPSFAIGNQLILKLYDDDPQDGWEAPEAPNESAIPSPALGFYPDSNPVRTEIAAISNVT 451

Query: 62   NRYKASINTGTVDPDEALPKLLADLKGAGWDKVQKRVQKQLDDF 105
             + +  ++  G VDP+E LP      L  AG  KV  E+Q+Q D++
Sbjct: 452 SEPSPALLKGAVDPEEYLPLPNDKLNEAGLQKVIDEMQRQFDEW 495
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2824

A DNA sequence (GASx1578R) was identified in *S.pyogenes* <SEQ ID 8147> which encodes the amino acid sequence <SEQ ID 8148>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> May be a lipoprotein

----- Final Results -----
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04515 GB:AP001509 unknown [Bacillus halodurans]
 Identities = 134/346 (38%), Positives = 205/346 (58%), Gaps = 10/346 (2%)

Query: 21   AACESKSASKDSDVKLLMVQVDKPNFDELMTTANKRIKEKTGATVDLQYIGWGDMDDK 80
            +A E+++   D  V L  Y +G    + + +M   N   +EK  ATVDL+ + WG++D++
Sbjct: 42   SANETATDLEH-VTLTWYMIGTPQDDLSLVMEEVNAYTEEKINATVDLRMLDWGEYDER 100

Query: 81   MSTIIASGENYDIAF----ANNYVVNAQKGAFADLTTLMPKYAKKTYKNLDPAYIKGNTI 136
            M  I  SGE YDIAF    ANNY +NA++GAF +L  L+ ++ ++  + +DPA+++G  +
Sbjct: 101  MQVITTSGEAYDIAFTSSWANNYALNARRGAFLELNDLLDEHGQEMKELIDPAFLEGAQV 160

Query: 137  DGKLYAFPVDANVYAQQMLSFNKELVDKYGLDISNIKSYADAENVLKQFHEKEPNTAAFA 196
             DGKLYA P +  V  Q +LSFN KLV+K+ LD+S++ S AD E +L    E+E +    A
Sbjct: 161  DGKLYAVPTNKEVQQQAVLSFNNELVEKHNLDLSSVHSLADLEPLLAVIKEEESDVTPIA 220

Query: 197  IGQVFSMSGDYDYPLTKTQPFAVKIDEGKPTIINQYEDESFKNNLRLMHKWYKEGLIPTD 256
              F     +D  L +  PFA +++      +IN+YE++      L+ MH +YK+G I  D
Sbjct: 221  ---TFDAYLPPDSILQEEMPFAPRLEGNTNSVINKYEEDITMETLKTMHDYYKEGYIRPD 277

Query: 257  AATWTEGYPLEGNTWPMRETQGPWDYGDTILTNAAGKDIVSRPLTKPLKTTSQAQMANF 316
             AAT+T+ +PLE   WF+R+E   P  Y + I T  AG +I +RPL +P     +
Sbjct: 278  AATSTDSWPLETPNWFVRKELYQP--YAELIWTRTAGYEIATRPLHEPYTFNNSVTGSMQ 335

Query: 317  VVSSVSKNKEKAVEVLSLLNSDPKLLNGLVYGVEGKAWEKIGDKKI 362
              +S+ SKN E+A+  L+LLNSDP L N L  G+EG  +E++ D  I
Sbjct: 336  AISATSKNPRAMMFLNLLNSDPYLRNLLDKGIEGVHYEKLEDGTI 381
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2825

A DNA sequence (GASx1582) was identified in *S.pyogenes* <SEQ ID 8149> which encodes the amino acid sequence <SEQ ID 8150>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0454(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2826

A DNA sequence (GASx1584R) was identified in *S.pyogenes* <SEQ ID 8151> which encodes the amino acid sequence <SEQ ID 8152>. Analysis of this protein sequence reveals the following:

```
Possible site: 41

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3105(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>

RGD motif: 3-5
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAG21428 GB:AF307332 meningioma-expressed antigen 5s splice
            variant [Homo sapiens]
 Identities = 94/271 (34%), Positives = 148/271 (53%), Gaps = 14/271 (5%)

Query: 120 GIIEGFYGTPWTRBERLDCLRPIGNKRMWTYMYAPKDDDYQRKLWRDLYPRDWVTYFKEL 179
           G++EGFYG PW  E+R +  R +     +NTY+YAPKDD    R  WR++Y  +       L
Sbjct: 63  GVVEGFYGRPWVMECRKELFRRLCKWELNTYLYAPKDDYKHRMFWRRMYSVEEAEQLMTL 122

Query: 180 LAVAKEEGLDFWYMISPGLDPDYTKRADYQLLYQKLQQLLALGVCHPGLLLDDIDYQIVD 239
           ++ A+E  ++F Y ISPGLD  ++    +   L +KL Q+   G   F LL DDID+ +
Sbjct: 123 ISAAREYEIEFIYAISPGLDITFSNPKEVSTLKRKLDQVSQFGCRSFALLPDDIDHNMCA 182

Query: 240 AVERRFKKTAYAQAHLATEVHHFLNQQHAAPELVICPTE------YDNHHDSTYLQELSE 293
           A +  F   A+AQ  +  E++ +L +        + CPTE      Y N   S YL+ + E
Sbjct: 183 ADKEVFSSFAHAQVSITNEIYQYLGEPET---FLFCPTEYCGTFCYPNVSQSPYLRTVGE 239

Query: 294 RIPKEVAFPWTGPSTLASQISQADIETMAAVYQRPIIIWDNIPVNDYQKDPERLFLTPFA 353
           ++   +   WIGP  ++ +I    IE ++ + +R  +IWDNI  NDY  D +RLFL P+
Sbjct: 240 KLLPGIEVLWTGPKVVSKEIPVESIEEVSKIIKRAPVIWDNIHANDY--DQKRLFLGPYK 297

Query: 354 NRSPFLCQPDYQVKGIVSNPMISWELSKLTL 384
           RS  L     ++KG+++NP   +E + + +
Sbjct: 298 GRSTELIP---KLKGVLTNPNCMPKANYVAI 325
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2827

A DNA sequence (GASx1585R) was identified in *S.pyogenes* <SEQ ID 8153> which encodes the amino acid sequence <SEQ ID 8154>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4469(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2828

A DNA sequence (GASx1587) was identified in *S.pyogenes* <SEQ ID 8155> which encodes the amino acid sequence <SEQ ID 8156>. Analysis of this protein sequence reveals the following:

```
Possible site: 47

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3082(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04509 GB:AP001509 unknown conserved protein in others
           [Bacillus halodurans]
 Identities = 221/425 (52%), Positives = 296/425 (69%), Gaps = 4/425 (0%)

Query: 12  RPIP*SVSQFMAKVESLCGDQHPDWALNPKTSPTNTLETTLKTYEDGTSPLLTGDIPAMW 71
           + IP S+   +A+V++   D       L F+  F NT  TT++  E GT F++TGDIPAMW
Sbjct: 4   KKIPRSLQAIAQVKAHYADDQELQTL-FEQCFLNTYLTTIQEDEQGT-FVVTGDIPAMW 61

Query: 72  LRDSTAQMKPYLPLAKEDEEIRKIIAGLVKRQFRYICIDPYANAFNEEANKKGHQTDMTQ 131
           LRDS+AQ++PYL + KED ++ ++I G+++RQ+RYI  DPYANAFN+ AN++GHQ D T+
Sbjct: 62  LRDSSAQVRPYLTVVKEDADMARMIKGVIERQWRYILHDPYANAFNQTANKQGHQQDRTE 121

Query: 132 MNFWIWERKYEIDCLCYPIQLAYLLYRETGSTDQFNDDFHRGVELILDLWTVEQDH-AQS 190
           M+P +WSRKYE+D LCYPIQLAYL ++ TG          + +E I   +W +EQDH A+S
Sbjct: 122 MSPLVWSRKYELDSLCYPIQLAYLYWKATGDDSVLQPTLKQVLETIYRIWKTEQDHEAKS 181

Query: 191 PYLPERDTWRKKDTLTHAGKGSPVAPTGMTWSGPRPSDDACQYGYLLPSNMPAVVVLSYL 250
            Y FERD  R  DTL   GKG    PTGMTWSGPRPSDDAC YGYLIP+NMPAVVV +Y
Sbjct: 182 SYSFERDDCRVSDTLLRKGKGGYSVPTGMTWSGPRPSDDACLYGYLIPANMPAVVVSNYA 241

Query: 251 EDLYNNLFHNEPVATRAKQLKEAIQGGIADHALVQNSKGETIYAYEVDGLQQFSIMDDAN 310
           +L   +     +A   ++L+  I+ GI  +  + +      IY YE DG G+ ++MDDAN
```

```
Sbjct: 242 VELLTAM-EEIKLAKEFRELEADIRQGIGQYGKMDHPVYGEIYVYETDGNGRVNLMDDAN 300

Query: 311 IPSLLAAPYLGFCTKDDPIYLAYKRTILSQENPVYYQGNAAAGIGSSHTPENYIWHIALA 370
           +PSLLA PYLG+ T DDP+Y  TRR ILS++NPYYY+G+ A G+GS HTP++Y+WHI+LA
Sbjct: 301 VPSLLAIPYLGYTTADDPVYQNTRKFILSRDNPYYYRGSYAKGVGSPHTPDHYVWHISLA 360

Query: 371 LQGLTALDQDSKKEMLDLLVATDAGTHLMHEGFDVNDPYQYTEEWFSWANMMFCELLDY 430
           +QG+TA+D   KK+++ +   T A T+ MHEGFDV+ P QYTR WF+WAN MF E LL
Sbjct: 361 IQGMTAIDSKEKKQIVAMFKQTHADTYFMHEGFDVDRPEQYTRSWFAWANSMFSEFLLSE 420

Query: 431 LGFSI 435
             G  +
Sbjct: 421 AGIYV 425
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2829

A DNA sequence (GASx1588) was identified in *S.pyogenes* <SEQ ID 8157> which encodes the amino acid sequence <SEQ ID 8158>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.5250(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB04508 GB:AP001509 unknown conserved protein in others
            (divided) [Bacillus halodurans]
 Identities = 312/737 (42%), Positives = 426/737 (57%), Gaps = 21/737 (2%)

Query: 123 FPDTFGNMGQPTQLMLKAGLQAAAPGRGIRPTGPFNNQVDTSEKYSSQPSEISWQGPDNSR 182
           FPDTFG  GQ PQL+ +AG++AA FGRG+ PTGPFNNQV   + YSS FSE+ W+ PD S+
Sbjct: 4   FPDTFGIYGQAPQLLAQAGIRAAVFGRGVIPTGPFNNQVQHDD-YSSPFSELIWEAPDGSQ 62

Query: 183 ILGLLFANWYSNGNEIPTTRAEARLFWDKKLADAERFASTKHLLMMNGCDHQPVQLDVTK 242
           ++G+L ANWYSNGNEIPT E EA+ FW KKL DAERFAST  LL MNGCDHQPVQ DVT+
Sbjct: 63  VIGILLANWYSNGNEIPTDEDEAQTFWVKKLRDAERFASTSQLLFMNGCDHQPVQKDVTQ 122

Query: 243 AIALANQLYPDYEFVHSCFEDYLADLADDLPENLSTVQGEITSQETDGWYTLANTASARI 302
           AI +A  L+PD  F HS F DYL  + ++LP+ L  + GE+ +Q+TDGW TL NTASARI
Sbjct: 123 AIKVAETLFPDVAFKHSNFHDYLTQIKKELPKELQKITGELRNQKTDGWSTLVNTASARI 182

Query: 303 YLKQANTRVSRQLRNITRPLAAMAYEVTSTYVPHDQLRYAWKTLMQNHPHDSICGCSVDSV 362
           YLKQAN R    L N+ EP+ +      +   D   Y WK LM+NHPHDSICGCS+D+V
Sbjct: 183 YLKQANDRCQTLLTNVLEPMCLLV--ENKSLHRDFSEYYWKLLMENHPHDSICGCSIDAV 240

Query: 363 HREMMTRFEKAYEVGHYLAKEAAKQIADAIDTRDFPMDSQPPVLFNTSGHSKTSVAELSL 422
           HREM TRFEK          E  K+IA  I+T      ++ P V+  T+G S   V    +
Sbjct: 241 HREMKTRFEKVEAGATTFIAEQGKKIAAQINTLHDSSEKAIPLVVLKTNGTSGKRVVRHKV 300

Query: 423 TWKKYHFQQRFPKEVYQEAQEYLARLSQSFQIIDTSGQVKPEAEILGTSIAFDYDLPKRS 482
            KK +F +      ++   + L  +      ++   +      E+    + F YDLP+
Sbjct: 301 AMKKIYFDEM----DPRHIPDRLKSIVMPTYRLEFPNKGSVPIEVQDAGVRPGYDLPRDG 356

Query: 483 FREPYFAIKVRLRLPITLDAMSWKTLALKLG------NETTPSETVSLYDDSNQCLENGF 536
           FR PY+A      L +T    S   L  + G      + T +     + D S    LEN
Sbjct: 357 FRRPYYA----RELEVTFSYDSDLYLGYRCGFLVPVSEKQTRARKELIGDPSMNTLENEA 412
```

```
Query: 537 LKVMIQTDGRLTTITDKQSGLIYQDLLRFEDCGDIGNEYISRQPNHDQPFYADQGTIKLNI 596
           +KVMI  +G  +I DK +G  Y+ L  +ED GDIGNEY+  +  +  +  + + I
Sbjct: 413 MKVMIHRNGSYSIILDKTTGFKYTRHIGIYEDVGDIGNEYMFKASSDGVRYTTSACEASIRI 472

Query: 597 ISNTAQVAELEIQQTFAIPISADKLLQAEMEAVIDITEQARRSQEKARLTLTTLIRMEK 656
           I N +  A +EI QT ++P +AD+ L+ E E ++    +R+A RS+E+ ++TL T + +E+
Sbjct: 473 IENNSLCATVEICQTLSVPAAADERLKEEQRRLVWHPDRKAGRSKERTDITLRTELTLEQ 532

Query: 657 NNPRLQFTTRPDNQMTNHRLRVLFPTHLKTDHHLADSIPETVKRPNHPDATFWKNPSNPQ 716
               L+      DN   +HR+R LFP       +H ADSI+E V+RPN PD   W+NP+
Sbjct: 533 GAKGLKVNVNIDNTAKDHRMRALFPVERARGNHYADSIYEIVERPNTPDPK-WQNPAPDH 591

Query: 717 HQECFVSLPDGENGVTIGNYGLNEYEILPDTNTIAITLLRSVGEMGDWGYFPTPEAQCLG 776
           H +  VSL +GE G+TI   GL+EYEI+ D  +LA+TLLRSVGE+GDWG F TPEAQC G
Sbjct: 592 HMQRLVSLDNGEYGLTIATKGLHEYEIVSD--SIAVTLLRSVGELGDWGLFETPEAQCFG 649

Query: 777 KHSLGYSFESITKQTQFAS-YWRAQEQQVFVITTQTWQHEGTLAAEYSYLTGTNDQVALT 835
           ++  +        A+ Y A +  V   QT Q  G L   +    + + LT
Sbjct: 650 QNEAQFVLLPHKGDVLSANVYVAAYDDPVEPTVIQTEQSMGPLPHATNLFQWSGEGLVLT 709

Query: 836 AFKRRLADNALITRSYN 852
           A K  +   +I R +N
Sbjct: 710 ACKPTMDGRGMILRWFN 726
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2830

A DNA sequence (GASx1589R) was identified in *S.pyogenes* <SEQ ID 8159> which encodes the amino acid sequence <SEQ ID 8160>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have a cleavable N-term signal seq.
    INTEGRAL     Likelihood =-11.30   Transmembrane  203 - 219 ( 195 - 221)
    INTEGRAL     Likelihood = -8.17   Transmembrane   61 -  77 (  59 -  82)
    INTEGRAL     Likelihood = -3.98   Transmembrane  107 - 123 ( 107 - 124)
    INTEGRAL     Likelihood = -3.40   Transmembrane   39 -  55 (  38 -  58)
    INTEGRAL     Likelihood = -2.34   Transmembrane  129 - 145 ( 126 - 145)
    INTEGRAL     Likelihood = -2.07   Transmembrane   89 - 105 (  87 - 105)

----- Final Results -----
                bacterial membrane --- Certainty=0.5522(Affirmative) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAC10175 GB:AJ278302 histidine kinase [Streptococcus pneumoniae]
 Identities = 114/432 (26%), Positives = 219/432 (50%), Gaps = 10/432 (2%)

Query: 21  LTLKLPSFVSAIPLRLKNIFYLSLSMVLPQVVFWAPFPDHFILDVVMLAQF---LPFALI 77
           L + +F  V  I L  + IF   L  +L  VVF       +++  V L+ F   L+   +
Sbjct: 16  LKIVIFFKVDGISLTFERIFKAFLFKILLAVVFGML---GYMVGNVYLSYFMEPLYGIGL 72

Query: 78  ALYYGKSIKAKFLMFYAFFPLVSISLVKRFIVFPVWPLFGMFYSVVKHNTLLIYSITCFS 137
           +    + +  K L+FY  FF++ ++L  R + +PV+P  G    V    + +  I  F+
Sbjct: 73  SFLLLRELPKKLLLFYGLFPWILVNLFYRGVSYFVLPFLQQG-QVYDDYSPIWLCIIIFN 131

Query: 138 IFLIYRCIQVFHFDPSTWRQYPQSHRASKLLVFTNSGMALYYLCVQGIDVMSPLSLGLAT 197
               F+    ++  +DF++ R+        K L   N  M  YYL +Q +        G+ +
Sbjct: 132 FFISLAFLKWLDYDFTSLRKGILDKDPQKSLTQINWIMGAYYLAVIQNLSYFEYE-QGIQS 190
```

```
Query: 198 TTARSIIVLFYFILMLTLLIHLSRYVKQNSIEAIVQQKR--YRELINYSCHLGLLYQDIQ 255
           TT R +I++FY + F+ ++  L+ Y+K    E + Q+++  YRE+  YS+H+  LY++++
Sbjct: 191 TTVRHLILVFYLLFFMGIIKKLDYYLKDKLHERINQEQDLRYREMSRYSRHIEELYKEVR 250

Query: 256 ELRRLLGTVSSRLKIGIEQNDISVRLTVEGIINARKNNAKDDRLDLTCLDKLQVEAIRH 315
             R   T + + L++GIE+ D+  ++  Y+ +L       +D++ DL  L  ++  A++
Sbjct: 251 SPRHDYTNLLTSLRLGIEEEDMEQIKEIYDSVLKDSSEKLQDNKYDLGRLVNVRDRALKS 310

Query: 316 IVLAKLIEAKNKKLKVEVSIPNCIATFFLEVVDPTKLLSFLLDNAIRMSLETKQPCLSIA 375
           ++  K I+A++K +   V +P  I    + ++DF  ++S L DNAIE S+E  QP +SIA
Sbjct: 311 LLAGKPIKARDKNIVFNVKVPREIQVEGVSLLDFLTVVSILCDNAIEASVKACQPHVSIA 370

Query: 376 FLDQNHKLVIVIQSSTKQGQDDSQSVFAIPALKKRDDWQFDLRNVTTILNRYDYLTISSQ 435
           F     +  +I++S K+   D   +F+  A  K ++    L  V  I+  +  ++++
Sbjct: 371 PFKNGAQETFIIENSIKEEGIDISEIPSFGASSKGEERGVGLYTVMKIVRSHPNTSLNTT 430

Query: 436 IHDGILTQLIEI 447
           D +  Q++ +
Sbjct: 431 CQDEVFRQVLTV 442
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2831

A DNA sequence (GASx1593R) was identified in *S.pyogenes* <SEQ ID 8161> which encodes the amino acid sequence <SEQ ID 8162>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -1.28   Transmembrane     2 -  18 (   1 -  18)

----- Final Results -----
                bacterial membrane --- Certainty=0.1510(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2832

A DNA sequence (GASx1594) was identified in *S.pyogenes* <SEQ ID 8163> which encodes the amino acid sequence <SEQ ID 8164>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -3.93   Transmembrane    76 -  92 (  76 -  92)

----- Final Results -----
                bacterial membrane --- Certainty=0.2572(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
                bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF61313 GB:U96166 unknown [Streptococcus cristatus]
 Identities = 31/66 (46%), Positives = 40/66 (59%), Gaps = 2/66 (3%)

Query: 14 LLGRILSKVVGRLTSCIENETTKIRNHSRQNDTIGLNHLLGNLKTVENPEIILKTINVYS 73
          + G  +SK    +   + E  K+ ++  ND IG N LLG+LKTVENPEII +   VYS
Sbjct: 30 VFGMDVSKTSSEVAILVNGE--KVHGYTILNDAIGFNRLLGDLKTVENPEIIPEATGVYS 87

Query: 74 RRLQVF 79
          RRLQ F
Sbjct: 88 RRLQAF 93
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2833

A DNA sequence (GASx1598) was identified in *S.pyogenes* <SEQ ID 8165> which encodes the amino acid sequence <SEQ ID 8166>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2117(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2834

A DNA sequence (GASx1608) was identified in *S.pyogenes* <SEQ ID 8167> which encodes the amino acid sequence <SEQ ID 8168>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2835

A DNA sequence (GASx1619) was identified in *S.pyogenes* <SEQ ID 8169> which encodes the amino acid sequence <SEQ ID 8170>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2916(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2836

A DNA sequence (GASx1621) was identified in *S.pyogenes* <SEQ ID 8171> which encodes the amino acid sequence <SEQ ID 8172>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1899(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
alpha subunit [Escherichia coli]
 Identities = 110/211 (52%), Positives = 153/211 (72%)

Query: 7   KEITIKEAVAHVKDGDTIMVGGFMTNGTPEKLIDALVEKGVKDLTLICNDAGFPDKGVGK 66
           K +T++ A      +DG TIMVGGFM  GTP +L++AL+E GV+DLTLI ND  F D G+G
Sbjct: 4   KLMTLQDATGFFRDGMTIMVGGFMSIGTPSRLVEALLESGVRDLTLIANDTAFVDTGIGP 63

Query: 67  MVANKQFSTIIASHIGLNREAGRQMTEGETVIDLVPQGTLAERIRSGGPGLGGFLTPTGI 126
           ++ N +    +IASHIG N E GR+M  GE  + LVPQGTL E+IR GG GLGGFLTPTG+
Sbjct: 64  LIVNGRVRKVIASHIGTNPETGRRMISGEMDVVLVPQGTLIEQIRCGGAGLGGFLTPTGV 123

Query: 127 GTEVAKGKEVITIDGKDYLLEKPLKADVALIFANKADKNGNLQYAGSENNFNPIVMAANAK 186
           GT V +GK+ +T+DGK +LLE+PL+AD+ALI A++ D  GNL Y  S  NFN ++A  A
Sbjct: 124 GTVVERGKQTLTLDGKTWLLSRPLRADLALIRAHRCDTLGNLTYQLSARNFNPLIAIAAD 183

Query: 187 TTIVEAREIVDVGQMDPNFVHTPGIFVNYLV 217
            T+VE  E+V+ G++ P+ + TPG  +++++
Sbjct: 184 TTLVEPDELVETGELQPDHIVTPGAVIDHII 214
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2837

A DNA sequence (GASx1622) was identified in *S.pyogenes* <SEQ ID 8173> which encodes the amino acid sequence <SEQ ID 8174>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4668(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAD54948 GB:AF157306 acetoacetate:butyrate/acetate coenzyme A
          transferase [Clostridium beijerinckii]
 Identities = 121/214 (56%), Positives = 161/214 (74%), Gaps = 5/214 (2%)

Query: 7   VLSKEEIQTRIAKRVAQELSHNTLVNLGIGLPTKVANYIPEGVTITLQSENGFVGLTGLT 66
           VL+KE I    AKRVA+EL+    LVNLGIGLPT VANY+P+ + IT +SENG VG+  +
Sbjct: 6   VLAKEII----AKKVAKELKXGQLVNLGIGLPTLVANYVPKEMNITFESENGMVGMAQMA 61

Query: 67  DD-HYDPTIVNAGGQPVSIAPGGAFFDSSTSFGIIRGGHVAATVLGALQVDKEASIANYL 125
               DP I+NAGG+ V++ P GAFFDSSTSF +IRGGHV   VLGALQVD+E ++AN++
Sbjct: 62  SSGENDPDIINAGGEYVTLLPQGAFFDSSTSFALIRGGHVDVAVLGALEVDEEGNLANWI 121

Query: 126 IPGKMVPGMGGAMDLLVGAKKVIVAMEHTMKGKAKILDKCTLPLTAQNVVNLIITEMGVF 185
           +P K+VPGMGGAMDL +GAKK+IVAM+HT KGK KI+ KCTLPLTA+  V+LI+TE+ V
Sbjct: 122 VPNKIVPGMGGAMDLAIGAKKIIVAMQHTGKGKPKIVKKCTLPLTAKAQVDLIVTELCVI 181

Query: 186 EYQDEGLCALEINPDYTFEDVQNVTEVTLLDKTN 219
           +  ++GL   EI+ D T ++++ +T+  LI   N
Sbjct: 182 DVTNDGLLFREIHKDTTIDEIKFLTDADLIIPDN 215
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2838

A DNA sequence (GASx1628R) was identified in *S.pyogenes* <SEQ ID 8175> which encodes the amino acid sequence <SEQ ID 8176>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1243(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2839

A DNA sequence (GASx1639R) was identified in *S.pyogenes* <SEQ ID 8177> which encodes the amino acid sequence <SEQ ID 8178>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -8.65   Transmembrane   55 -  71 (  44 -  73)
   INTEGRAL    Likelihood = -7.64   Transmembrane   13 -  29 (   5 -  31)

----- Final Results -----
              bacterial membrane --- Certainty=0.4461(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2840

A DNA sequence (GASx1643) was identified in *S.pyogenes* <SEQ ID 8179> which encodes the amino acid sequence <SEQ ID 8180>. Analysis of this protein sequence reveals the following:

```
Possible site: 35

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.0766(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2841

A DNA sequence (GASx1645R) was identified in *S.pyogenes* <SEQ ID 8181> which encodes the amino acid sequence <SEQ ID 8182>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2842

A DNA sequence (GASx1649R) was identified in *S.pyogenes* <SEQ ID 8183> which encodes the amino acid sequence <SEQ ID 8184>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0931(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2843

A DNA sequence (GASx1650) was identified in *S.pyogenes* <SEQ ID 8185> which encodes the amino acid sequence <SEQ ID 8186>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.5678(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2844

A DNA sequence (GASx1651R) was identified in *S.pyogenes* <SEQ ID 8187> which encodes the amino acid sequence <SEQ ID 8188>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2761(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2845

A DNA sequence (GASx1667R) was identified in *S.pyogenes* <SEQ ID 8189> which encodes the amino acid sequence <SEQ ID 8190>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2967(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2846

A DNA sequence (GASx1672) was identified in *S.pyogenes* <SEQ ID 8191> which encodes the amino acid sequence <SEQ ID 8192>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -3.82   Transmembrane    3 -  19 (   1 -  20)

----- Final Results -----
            bacterial membrane --- Certainty=0.2529(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2847

A DNA sequence (GASx1673R) was identified in *S.pyogenes* <SEQ ID 8193> which encodes the amino acid sequence <SEQ ID 8194>. Analysis of this protein sequence reveals the following:

```
Possible site: 38

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -8.86   Transmembrane   51 -  67 (  47 -  75)
   INTEGRAL    Likelihood = -5.20   Transmembrane   27 -  43 (  24 -  45)
   INTEGRAL    Likelihood = -3.66   Transmembrane  112 - 128 ( 112 - 131)

----- Final Results -----
```

```
bacterial membrane --- Certainty=0.4545(Affirmative) < succ>
 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF41294 GB:AE002440 conserved hypothetical protein [Neisseria
          meningitidis MC58]
 Identities = 61/148 (41%), Positives = 96/148 (64%)

Query: 1   LKKSITNEKALLAQSQEPGAQNTKPLTLLHIMIYVFAVIEALLKQIKFDGISFLGLLLM 60
           L  SI +EKA++A+G +++G  N+  L  +H + Y+   +   L    F+GIS +G L +
Sbjct: 19  LAVSIKHEKALIAKGAKQYGKTNSTLLAAVHTLYYLACPVWVWLSDTAPNGISLIGTLTV 78

Query: 61  LLSVAVLYEVTRILSDIWTVKLMLAKDHKYVDHWLFKTIKHPNYFLNIAPELVGIALLCH 120
           + S  +L  + + LS+IWTVK+ +  +H+     WLFKI +HPNYFLNI PELG+IALLC
Sbjct: 79  MASFVILSLIIKQLSEIWTVKIYILPNHQINRSWLFKTFRHPNYFLNIIPELIGIALLCQ 138

Query: 121 AKITAMLLFPCYIVVIYLRIREENKLLA 148
           A    ++  P Y++V++ RIR+E + +A
Sbjct: 139 AWYVLLIGLPIYLLVLFKRIRQEEQAMA 166
```

A related GBS gene <SEQ ID 9009> and protein <SEQ ID 9010> were also identified. Analysis of this protein sequence reveals the following:

```
Lipop: Possible site: -1   Crend: 0
McG: Discrim Score:      5.86
GvH: Signal Score (-7.5): 0.14
     Possible site: 60
>>> Seems to have a cleavable N-term signal seq.
ALOM program   count: 2 value:  -8.23 threshold:  0.0
   INTEGRAL    Likelihood = -8.23   Transmembrane   69 -  85 (  64 -  89)
   INTEGRAL    Likelihood = -3.29   Transmembrane  142 - 158 ( 140 - 159)
   PERIPHERAL  Likelihood =  1.70      123
 modified ALOM score:   2.15

*** Reasoning Step: 3

----- Final Results -----
               bacterial membrane --- Certainty=0.4291(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

The protein has homology with the following sequences in the databases:

```
42.1/64.0% over 168aa                                                 imported
  EGAD|177248| conserved hypothetical protein {Neisseria meningitidis} Insert characterized
  GP|7379797|emb|CAB84365.1||AL162755  putative  integral  membrane  protein  {Neisseria
meningitidis} Insert characterized
  GP|7226121|gb|AAF41294.1||AE002440 conserved hypothetical protein {Neisseria meningitidis
MC58} Insert characterized
  PIR|F81147|F81147 probable integral membrane protein NMA1102 - Neisseria meningitidis
(group B strain MD58, group A strain Z2491) Insert
 characterized

ORF00432(301 - 807 of 1140)
EGAD|177248|NMB0883(1  -  169  of  169)  conserved  hypothetical  protein  {Neisseria
meningitidis|GP|7379797|emb|CAB84365.1||AL162755  putative  integral  membrane  protein
{Neisseria  meningitidis|GP|7226121|gb|AAF41294.1||AE002440  conserved  hypothetical  protein
{Neisseria meningitidis MC58}PIR|F81147|F81147 probable integral membrane protein NMA1102
{imported] - Neisseria meningitidis (group B strain MD58, group A strain Z2491)
%Match = 19.0
%Identity = 42.0  %Similarity = 63.9
Matches = 71  Mismatches = 61  Conservative Sub.s = 37
```

```
237       267       297       327       357       387       417       447
SSGEYHLLTSDHSLV*IGKAXX*LIXXKEFTMSIIIGLMAAMFIIRLAYLKLSIANEKALRKNGAKEYGVGVSKAITVLH
                                  |::|: :::  |||| :| :|| :|||   |||:||   |  : :|
                                 WTMILSILSLFFIIRLLFLAVSIKHEKALIAKGAKQYGKTNSTLAAVH
                                           10        20        30        40

477       507       537       567       597       627       657       687
IIIYFSSVTEAILTKASFNFVSVIGLSLMIFSVFMLHTVTRLLGRIWTVKLMVDKNHQFVDHWLFRVVKHPNYFLNIAPE
 : |::      |:  :|| :|:||   :: |  :|  : : || |||||: :  |||    |||:  :||||||| ||
TLYYIACFVWVWLSDIAFNGISLIGTLTVMASFVILSLIIKQIGEIWTVKIYILPNHQINRSWLFKTFRHFNYFLNIIPE
         60        70        80        90       100       110       120

717       747       777       807       837       867       897       927
LLGVTLLCHAKYTALFVLPIYAFVIYLRIRESNLLLKTIIIPNGIKKSRVY*E*DK**T*KSFFVILSQ*EEVFISCFFS
|:|: |||:|  |  |  |||| :|:: |||:|   : |:
LIGIALLCQAWYVLLIGLPIYLLVLFKRIRQEEQAMATLF
         140       150       160
```

Based on this analysis, it was predicted that these proteins and their epitopes could be useful antigens for vaccines or diagnostics.

Example 2848

A DNA sequence (GASx1674R) was identified in *S.pyogenes* <SEQ ID 8195> which encodes the amino acid sequence <SEQ ID 8196>. Analysis of this protein sequence reveals the following:

```
Possible site: 23

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.3098(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2849

A DNA sequence (GASx1677R) was identified in *S.pyogenes* <SEQ ID 8197> which encodes the amino acid sequence <SEQ ID 8198>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -8.86   Transmembrane  254 - 270 ( 248 - 280)
   INTEGRAL    Likelihood = -7.01   Transmembrane  303 - 319 ( 296 - 322)
   INTEGRAL    Likelihood = -2.39   Transmembrane   74 -  90 (  74 -  91)
   INTEGRAL    Likelihood = -1.91   Transmembrane  201 - 217 ( 199 - 217)
   INTEGRAL    Likelihood = -1.91   Transmembrane  223 - 239 ( 220 - 240)
   INTEGRAL    Likelihood = -1.65   Transmembrane  118 - 134 ( 115 - 135)
   INTEGRAL    Likelihood = -1.49   Transmembrane   56 -  72 (  55 -  72)
   INTEGRAL    Likelihood = -0.32   Transmembrane   13 -  29 (  13 -  30)

----- Final Results -----
                bacterial membrane --- Certainty=0.4545(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB05126 GB:AP001511 unknown conserved protein [Bacillus halodurans]
 Identities = 249/534 (46%), Positives = 360/534 (70%)

Query: 12  QDIAFNFFGGLGLFLFSIKYMKDGLQQRAGDKLRYYIDKYTSNPFGILVGIAMSALIQS 71
           Q + F FFGGLG+FLF IKYMKDGLQ+ AG++LR  +DK+T+NF  G+L GI ++ L+Q+
Sbjct: 6   QTLLFMFFGGLGIFLFGIKYMGDGLQKRVAGKRLRDLLDKPTTNPLMGVLAGIVVTVLLQT 65

Query: 72  SSGVTVITVGLVSAGLLNLRQAIGIVMGANIGTTITSFLIGFKLGDYALMIFIGAACLF 131
           S+G TV+T+GLV+AG + L+QAIG++MGANIGTT+T+F+IG K+ +YALP+I +GAA +F
Sbjct: 66  STGTTVLTIGLVNAGFMTLKQAIGVIMGANIGTTVTAFIIGIKISEYALPIIAVGAALIF 125

Query: 132 FTSNKKLNNFGRIIFGVGGIFFSLNLMGDAMDPLKSVSAFQNYLATLGDKPPQGVFIGTA 191
           F  NKK+NN G++IFG G +F+ LN MG+ ++PL+ + AF +   ++ + P  GV IGT
Sbjct: 126 FIKNKKVNNIGQVIFGFGTLFYGLNTMGEGLNPLKELQAFADLTVSMSENPLLGVLIGTI 185

Query: 192 LTMLIQSSAAIIGILQGLFSGGLLTLQGAIPILLGSNIGTCITAVLAAIGSNIAAKRVAA 251
            T  +QSS+A IG+IQ L+  G + L  A+P+L G NIGT ITAVLAAIG+++AARR A
Sbjct: 186 FTAAVQSSSASIGLLQQLYDQGAMDLFAALPVLFGDNIGTTITNVLAALGASVAARRAAL 245

Query: 252 AHVLFNLIGTIIFMILVPFTSLMLWLGSKLSLTPEMTIAFSHGSFNITNTILLLIPFISL 311
            HV+FNLIGTII +II++PFT  + +L    +L   MTIAF+HG FN++NTI+  PFI +
Sbjct: 246 THVIFNLIGTIIVLIIIIPFTHFIAYLAEVFALNRPMTIAFAHGIFNVSNTIIQFPFIGI 305

Query: 312 LAMIVTRLIPGEDEVVKYEALYLDRILLITQAPSIALGNAHKELVHLASYAIQAFEASYSY 371
            LA+IVT+L+PG+D  ++Y+A +LD   +  +P+IALG A +E++ +A ++ +       Y
Sbjct: 306 LAIIVTKLVPGDDPYIEYKAKHLDPRFVGSSPAIALQQAKQEVLRMAEFSEKGLLEVSKY 365

Query: 372 IMTADGKFGEKVKRYERAVDTYDELTTYLVDISNEALSPSENEVLAGILDSSRDLERIG 431
           +    K  E   ++E A++ +D ++T YL+ IS+ +LS  +++    ++D+ RD+ERIG
Sbjct: 366 MENGQKKHAEMAVQFEDAINNLDRKITEYLISISSRSLSAQDSEMHGMLMDTVRDIERIG 425

Query: 432 DHSESLGILIEGIISKQIGFSISAPQELTEMYQLTHCLTLDAIRAIVDSDTDLAQTIVTR 491
           DH E++  L +   + ++  S  A  +L E++ LTH     +AI ++   D + A++++ +
Sbjct: 426 DHIENIVELKDYQKANKVKISEKALHDLQEMPDLITHSTLTRAIMSLETGDLEAARSVIEK 485

Query: 492 HKEIEEKERRLRKTHIKRLNCGECTAQAGINFIDIISHYTRITDHALNLAEKVL 545
            + I++ ER+LRK HI R+N G CT  AGI F+DI+S+  RI DH++N+AE V+
Sbjct: 486 EEHIDQMERKLRKQHIIRVNEGNCTGAAGIVFVDIVSNLERIGDHSVNIAEAVI 539
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2850

A DNA sequence (GASx1678R) was identified in *S.pyogenes* <SEQ ID 8199> which encodes the amino acid sequence <SEQ ID 8200>. Analysis of this protein sequence reveals the following:

```
Possible site: 48

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2940(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2851

A DNA sequence (GASx1685R) was identified in *S.pyogenes* <SEQ ID 8201> which encodes the amino acid sequence <SEQ ID 8202>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL     Likelihood = -7.11   Transmembrane   13 -  29 (   9 -  31)

----- Final Results -----
               bacterial membrane --- Certainty=0.3845(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2852

A DNA sequence (GASx1695R) was identified in *S.pyogenes* <SEQ ID 8203> which encodes the amino acid sequence <SEQ ID 8204>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1357(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2853

A DNA sequence (GASx1698) was identified in *S.pyogenes* <SEQ ID 8205> which encodes the amino acid sequence <SEQ ID 8206>. Analysis of this protein sequence reveals the following:

```
Possible site: 33

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1970(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2854

A DNA sequence (GASx1713) was identified in *S.pyogenes* <SEQ ID 8207> which encodes the amino acid sequence <SEQ ID 8208>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3092(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2855

A DNA sequence (GASx1737) was identified in *S.pyogenes* <SEQ ID 8209> which encodes the amino acid sequence <SEQ ID 8210>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1878(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2856

A DNA sequence (GASx1748R) was identified in *S.pyogenes* <SEQ ID 8211> which encodes the amino acid sequence <SEQ ID 8212>. Analysis of this protein sequence reveals the following:

```
Possible site: 23

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2841(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2857

A DNA sequence (GASx1750R) was identified in *S.pyogenes* <SEQ ID 8213> which encodes the amino acid sequence <SEQ ID 8214>. Analysis of this protein sequence reveals the following:

```
Possible site: 59

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.22   Transmembrane   18 -  34 (  18 -  34)

----- Final Results -----
               bacterial membrane --- Certainty=0.1489(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2858

A DNA sequence (GASx1754) was identified in *S.pyogenes* <SEQ ID 8215> which encodes the amino acid sequence <SEQ ID 8216>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2859

A DNA sequence (GASx1759) was identified in *S.pyogenes* <SEQ ID 8217> which encodes the amino acid sequence <SEQ ID 8218>. Analysis of this protein sequence reveals the following:

```
Possible site: 36

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1534(Affirmative) < succ>
```

```
bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2860

A DNA sequence (GASx1764R) was identified in *S.pyogenes* <SEQ ID 8219> which encodes the amino acid sequence <SEQ ID 8220>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -6.74   Transmembrane   90 - 106 (  87 - 121)
   INTEGRAL    Likelihood = -4.57   Transmembrane  210 - 226 ( 205 - 229)
   INTEGRAL    Likelihood = -4.19   Transmembrane   43 -  59 (  42 -  62)
   INTEGRAL    Likelihood = -3.77   Transmembrane  137 - 153 ( 137 - 155)

----- Final Results -----
              bacterial membrane --- Certainty=0.3697(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2861

A DNA sequence (GASx1768R) was identified in *S.pyogenes* <SEQ ID 8221> which encodes the amino acid sequence <SEQ ID 8222>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-12.37   Transmembrane   26 -  42 (  17 -  47)
   INTEGRAL    Likelihood = -7.54   Transmembrane   53 -  69 (  46 -  73)
   INTEGRAL    Likelihood = -3.29   Transmembrane  209 - 225 ( 209 - 225)
   INTEGRAL    Likelihood = -2.13   Transmembrane   82 -  98 (  82 -  98)
   INTEGRAL    Likelihood = -1.65   Transmembrane    9 -  25 (   9 -  25)
   INTEGRAL    Likelihood = -0.85   Transmembrane  117 - 133 ( 117 - 134)

----- Final Results -----
              bacterial membrane --- Certainty=0.5946(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB84959 GB:AE000829 conserved protein [Methanobacterium
        thermoautotrophicum]
 Identities = 54/192 (28%), Positives = 90/192 (46%), Gaps = 6/192 (3%)
```

```
Query: 7    TKLLLLVLANACPFFRVDGFLEFIIVIFLLLLLSALNKKKLA--FKLAVVYLLMIGLSVI 64
             +KL ++V A      F  D  L  I+ +     L++      + A  F   ++   ++ L++I
Sbjct: 32   SKLTVVVSATLLSTFISDLTLLIIMGVIFTALIAHSGSLRFAAPFLSFIILFWLVSLAII 91

Query: 65   PLSIFPSYLDILLSFVSIAGRLVFPSLLAGLITIKTFTTIYELVHGLRKWRFDEVWLLTLA 124
              +   S   H + F+S+      F   AGL    TT   +L   LR  R P    + TL
Sbjct: 92   MVL---SGNPHTWGFLSLFFARFFIISAAGLSFAPTTEPQKLAESLRSVRIPGEIVFTLT 148

Query: 125  VMCRFIPMIRQECCVIHRSLKIRGIIITKWSILIRPKQYLEYLMVPLLLSLIRSSQELTI 184
             V  R+IP +  E   I  SLK+R   L+  SI+ KP      L++P+++  ++ S E+ I
Sbjct: 149  VALRYIPALAVEASSIWDSLKLR-TSLSGSSSIIRRPSLLYRGLIIPMIIRTVKISDEVAI 207

Query: 185  ASLTKGLAVNKG 196
             A+ T+G    +G
Sbjct: 208  AAETRGFNPREG 219
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2862

A DNA sequence (GASx1769R) was identified in *S.pyogenes* <SEQ ID 8223> which encodes the amino acid sequence <SEQ ID 8224>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -7.32   Transmembrane  164 - 180 ( 158 - 186)
   INTEGRAL    Likelihood = -4.67   Transmembrane   85 - 101 (  84 - 105)
   INTEGRAL    Likelihood = -3.03   Transmembrane   42 -  58 (  42 -  61)
   INTEGRAL    Likelihood = -2.76   Transmembrane  118 - 134 ( 117 - 134)
   INTEGRAL    Likelihood = -2.07   Transmembrane   64 -  80 (  64 -  82)
   INTEGRAL    Likelihood = -1.22   Transmembrane   18 -  34 (  17 -  34)

----- Final Results -----
               bacterial membrane --- Certainty=0.3930(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2863

A DNA sequence (GASx1776R) was identified in *S.pyogenes* <SEQ ID 8225> which encodes the amino acid sequence <SEQ ID 8226>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -6.37   Transmembrane    4 -  20 (   1 -  22)
   INTEGRAL    Likelihood = -0.43   Transmembrane  261 - 277 ( 261 - 278)

----- Final Results -----
               bacterial membrane --- Certainty=0.3548(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2864

A DNA sequence (GASx1777R) was identified in *S.pyogenes* <SEQ ID 8227> which encodes the amino acid sequence <SEQ ID 8228>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -8.17   Transmembrane 1217 -1233 (1215 -1235)

----- Final Results -----
               bacterial membrane --- Certainty=0.4270(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF53254 GB:AE003639 CG16974 gene product [Drosophila
          melanogaster]
 Identities = 84/238 (35%), Positives = 133/238 (55%), Gaps = 10/238 (4%)

Query: 516 LRLDHYELTDISLL--KHAKNITELHLDGNQITEIPKELFSQMKQLRFLNLRSNHLTYLD 573
           L +    L++ SLL  ++ K + E+HLD +++T +P+    ++ +LR LNL  N LT L
Sbjct: 232 LEMSGNRLSNCSLLNLQYMKQLQELHLDRSELTYLPQRFLGELSELRMLNLSQNLLTELP 291

Query: 574 KDTFKSNAQLRELYLSSNFIHSLEGGLFQSLHHLEQLDLSKNRIGRLCDNPFEGLSRLTS 633
           +D F    +L  LYLS N +  L   LFQ+   L+ LDLS NR+    DN F    +L
Sbjct: 292 RDIFVGALKLERLYLSGNRLSVLPPMLFQTAADLQVLDLSDNRLLSFPDNFFARNGQLRQ 351

Query: 634 LGFAENSLEEIPEKALEPLTSLNFIDLSQNNLALLP-KTIEKLRALSTIVASRNHITRID 692
           L    N L+ I + +L  L  L  +DLSQN+L+++  K  E L  L  +  S N++T +
Sbjct: 352 LHLQRNQLKSIGKHSLYSLRELRQLDLSQNSLSVIDRKAFESLDHLLALNVSGNNLTTLS 411

Query: 693 NISPKNLPKLSVLDLSTNEISNLPNGIFKQNNQL-------TKLDFFNNLLTQVEESV 743
           +I F++L  L  LDLS N+   LP+G+F++    L       T ++ F+N +++ +ES+
Sbjct: 412 SIIPQSLHALRQLDLSRNQFKQLPSGLFQRQRSLVLLRIDETPIEQFSNWISKYDESL 469
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2865

A DNA sequence (GASx1778R) was identified in *S.pyogenes* <SEQ ID 8229> which encodes the amino acid sequence <SEQ ID 8230>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1067(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2866

A DNA sequence (GASx1779) was identified in *S.pyogenes* <SEQ ID 8231> which encodes the amino acid sequence <SEQ ID 8232>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1885(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2867

A DNA sequence (GASx1786R) was identified in *S.pyogenes* <SEQ ID 8233> which encodes the amino acid sequence <SEQ ID 8234>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0612(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2868

A DNA sequence (GASx1790) was identified in *S.pyogenes* <SEQ ID 8235> which encodes the amino acid sequence <SEQ ID 8236>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
              bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2869

A DNA sequence (GASx1791R) was identified in *S.pyogenes* <SEQ ID 8237> which encodes the amino acid sequence <SEQ ID 8238>. Analysis of this protein sequence reveals the following:

```
Possible site: 43

>>> Seems to have no N-terminal signal sequence
   INTEGRAL     Likelihood = -0.90   Transmembrane   28 -  44 (  28 -  44)

----- Final Results -----
               bacterial membrane --- Certainty=0.1362(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

A related sequence was also identified in GAS <SEQ ID 9155> which encodes the amino acid sequence <SEQ ID 9156>. Analysis of this protein sequence reveals the following:

```
     Possible site: 25
>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty= 0.300(Affirmative) < succ>
               bacterial membrane --- Certainty= 0.000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty= 0.000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA24923 GB:L06331 endoglycosidase [Chryseobacterium
           meningosepticum]
 Identities = 105/322 (32%), Positives = 153/322 (46%), Gaps = 53/322 (16%)

Query: 106 ADKQAQELAKMKIPEKIPMKPLHGSLYGGYFRTWHDKTSDPTEKDKVNSMGELPKEVDLA 165
           A K    ++ + +   I  K     +  GY+RTW D     T    + SM  LP  +D+
Sbjct: 37  AQKSGVTVSAVNLSNLIAYKNSDHQISAGYYRTWRDSA---TASGNLPSMRWLPDSLDMV 93

Query: 166 PIPHDWTKDYSLPWKELATKHVPKLNKQGTRVIRTIPWRFLAGGDNSGIAEDTSKYPNTP 225
            +P D+T   + +W  L T +VP L+K+GT+VI T+      G  NS     T+
Sbjct: 94  MVFPDYTPPENAYWNTLKTNYVPYLHKRGTKVIITL------GDLNSA----TTTGQDS 143

Query: 226 EGNKALAKAIVDEYVYKYNLDGLDVDVEHDSIPKVDKKEDTAGVSRSIQVFEEIGKLIGP 285
            G  + AK I D++V +YNLDG+D+D+D+E              A + + +    + K  GP
Sbjct: 144 IGYSSWAKGIYDKWVGEYNLDGIDIDIE--------SSPSGATLTKFVAATKALSKYFGP 195

Query: 286 KGVDKSRLFIMDSTYMADKNP--LIERGAPYINLLLVQVYGSQGEKGGWEPVSNRPEKTM 343
           K     + F+ D+    ++NP      + AP  N + +Q YG              R    +
Sbjct: 196 KS-GTGKTFVYDTT----NQMPTNFFIQTAPRYNYVFLQAYG------------RSTTNL 237

Query: 344 KERWQQSYSKYIRPEQYMIGPSPYEENNAQEGNLWYDINSRKDRDKANGINTDITGTRAERY 403
                Y+ YI  +Q++ GPSPYEENN   GN W D+    +      NG     TG RA  Y
Sbjct: 238 TTVSGLYAPYISMKQFLPGPSPYEENNGYPGNYWNDVRYPQ-----NG-----TG-RAYDY 286

Query: 404 ARWQPKTGGVKGGIPSYAIDRD 425
           ARWQP T G KGG+FSYAI+RD
Sbjct: 287 ARWQPATGKKGGVFSYAIERD 307
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2870

A DNA sequence (GASx1803) was identified in *S.pyogenes* <SEQ ID 8239> which encodes the amino acid sequence <SEQ ID 8240>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2099(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2871

A DNA sequence (GASx1806R) was identified in *S.pyogenes* <SEQ ID 8241> which encodes the amino acid sequence <SEQ ID 8242>. Analysis of this protein sequence reveals the following:

```
Possible site: 54

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2706(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB16126 GB:Z99124 ribosomal protein S18 [Bacillus subtilis]
 Identities = 51/77 (66%), Positives = 63/77 (81%)

Query: 1  MAQQRRGGFKKRKKVDFIAANKIEYVDYKDTELLSRFVSERGKILPRRVTGTSAKNQRKV 60
          MA  RRGG  +R+KV +  +N I ++DYKD +LL +FVSERGKILPRRVTGT+AK QRK+
Sbjct: 3  MAGSRRGGRAKRRKVCYFTSNGITHIDYKDVDLLKKFVSERGKILPRRVTGTNAKYQRKL 62

Query: 61 TTAIKRARVMALMPYVN 77
          T AIKRAR MAL+PYV+
Sbjct: 63 TAAIKRARQMALLPYVS 79
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2872

A DNA sequence (GASx1809R) was identified in *S.pyogenes* <SEQ ID 8243> which encodes the amino acid sequence <SEQ ID 8244>. Analysis of this protein sequence reveals the following:

```
Possible site: 60
```

```
>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -7.59   Transmembrane   70 -  86 (  66 -  92)
   INTEGRAL    Likelihood = -6.42   Transmembrane   13 -  29 (   8 -  33)
   INTEGRAL    Likelihood = -5.68   Transmembrane   48 -  64 (  43 -  69)

----- Final Results -----
               bacterial membrane --- Certainty=0.4036(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2873

A DNA sequence (GASx1813R) was identified in *S.pyogenes* <SEQ ID 8245> which encodes the amino acid sequence <SEQ ID 8246>. Analysis of this protein sequence reveals the following:

```
Possible site: 56

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-10.51   Transmembrane  127 - 143 ( 113 - 147)
   INTEGRAL    Likelihood =-10.46   Transmembrane  151 - 167 ( 149 - 167)
   INTEGRAL    Likelihood = -4.41   Transmembrane   59 -  75 (  57 -  77)

----- Final Results -----
               bacterial membrane --- Certainty=0.5203(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB98363 GB:U67490 lipoprotein B (lppB) [Methanococcus
          jannaschii]
 Identities = 43/143 (30%), Positives = 68/143 (47%), Gaps = 7/143 (4%)

Query: 25  LLNVLLKIITGVMY--ILYPSLIPTLWQGMTFQLWLRLLIIPAVGFIALSYIRKRPDPP 82
           + + ++ II+   Y  I   S +IF   +    +L   L +   + F +L Y+       P
Sbjct: 181 IFDAIMPIISKTAYPLIAITSLIIPIKNRKFGMKLIPALFLAFMIAF-SLKYLVNE---P 236

Query: 83  RPYEKWNIKPLIDKDTKGRSMPSRHVPSATMISMCLLRYVYFGIVCLILSALLAICKVI 142
           RPY   +  L+  +    S PS H   A  ++  LL Y    GI+ L  + ++A  KV
Sbjct: 237 RPYLVLDNVHLLCNEGNEPSFPSGHTTLAFTLATSLLFYSKKLGILFLSWAIIVAYSRVY 296

Query: 143 AGIHYPKDVIVGYLIGLMLGLCL 165
            G+HYP DV+ G +IG+  G CL
Sbjct: 297 VGVHYPLDVLAGMIIGIPCG-CL 318
```

A related GBS gene <SEQ ID 9011> and protein <SEQ ID 9012> were also identified. Analysis of this protein sequence reveals the following:

```
Lipop: Possible site: -1   Crend: 9
McG: Discrim Score:      3.19
GvH: Signal Score (-7.5): -2.18
     Possible site: 55
>>> Seems to have a cleavable N-term signal seq.
ALOM program   count: 3 value: -11.78 threshold:  0.0
```

```
   INTEGRAL    Likelihood =-11.78   Transmembrane  126 - 142 ( 112 - 147)
   INTEGRAL    Likelihood =-11.30   Transmembrane  150 - 166 ( 147 - 166)
   INTEGRAL    Likelihood = -4.41   Transmembrane   58 -  74 (  56 -  76)
   PERIPHERAL  Likelihood =  3.29     107
modified ALOM score:   2.86

*** Reasoning Step: 3

----- Final Results -----
                bacterial membrane --- Certainty=0.5713(Affirmative) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

The protein has homology with the following sequences in the databases:

```
ORF01020(472 - 792 of 1098)
EGAD|44548|MJ0374(213  -  318  of  330)  conserved  hypothetical  protein  {Methanococcus
jannaschii} OMNI|MJ0374 conserved hypothetical protein SP|Q57819|Y374_METJA HYPOTHETICAL
PROTEIN  MJ0374.  GP|1591081|gb|AAB98363.1||U67490  lipoprotein  B  (lppB)  {Methanococcus
jannaschii} PIR|F64346|F64346 hypothetical protein MJ0374 - Methanococcus jannaschii
%Match = 6.8
%Identity = 30.8  %Similarity = 53.3
Matches = 33  Mismatches = 49  Conservative Sub.s = 24

222       252       282       312       342       372       402       432
EGVTKYLRRNKHVKHFAYAPQNAGGGATTVTLG*IMESYEQFYAKLSQPFRKSPQLIILLNFLLKIVTGMKYILYPSFL

VIAWLSGIFEMHKLLFTVGTIIGKLPRFLAVAYFGDVLGNINRLSDINIYLPYLINSHYNYIFDAIMPIISKTAYPLIAI
          130       140       150       160       170       180       190

462       492       522       552       582       612       642       672
IFTLWQGMTFQLWLRLLIIPAVGFIALSYIKKRLDFPPVEKWNIKPLIYKDTEGRSMDSKHVFSATMISMCLLRYTVYF
             ::|:    : |:    ::  :: ||||   :   |:  :    | || |   |  ::   || |   :
TSLIIFIKNRKFGMKLIFALPLAPMIAPSLAYLVNEPRPYLVLDNVHLLCNEGNEPSFPSGHTTLAFTLATSLLFYSKKL
          210       220       230       240       250       260       270

702       732       762       792       822       852       882       912
GIVCLILSVLLAICRVIAGIHYPKDVIVGYLIGLILGLCLFI*RVRSK*PQKQLDSCTIGLSLR*NGEKRWH*K*QMLHL
||: |  ::::|  ||  |:||| ||: | :||:  | ||
GILFLSWAIIVAYSRVYVGVHYPLDVLAGMIIGIFCG-CL/TRIDIYKLIDNI
          290       300       310       320       330
```

Based on this analysis, it was predicted that these proteins and their epitopes could be useful antigens for vaccines or diagnostics.

Example 2874

A DNA sequence (GASx1815R) was identified in *S.pyogenes* <SEQ ID 8247> which encodes the amino acid sequence <SEQ ID 8248>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.0888(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2875

A DNA sequence (GASx1825R) was identified in *S.pyogenes* <SEQ ID 8249> which encodes the amino acid sequence <SEQ ID 8250>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.16   Transmembrane     7 -  23 (   7 -  23)

----- Final Results -----
               bacterial membrane --- Certainty=0.1065(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2876

A DNA sequence (GASx1832) was identified in *S.pyogenes* <SEQ ID 8251> which encodes the amino acid sequence <SEQ ID 8252>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.0918(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2877

A DNA sequence (GASx1836R) was identified in *S.pyogenes* <SEQ ID 8253> which encodes the amino acid sequence <SEQ ID 8254>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4084(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2878

A DNA sequence (GASx1864R) was identified in *S.pyogenes* <SEQ ID 8255> which encodes the amino acid sequence <SEQ ID 8256>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.5280(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC36810 GB:L12244 ribosomal protein L28 [Bacillus subtilis]
 Identities = 45/62 (72%), Positives = 52/62 (83%)

Query: 1  MAKVCYFTGRKTVSGNNRSHAMNQTKRTVKPNLQKVTILVDGKPKKVWASARALKSGKVE 60
          MA+ C  TG+KT +GNNRSHAMN +KRT   NLQKV ILV+GKPKKV+ SARALKSGKVE
Sbjct: 1  MARKCVITGKKTTAGNNRSHAMNASKRTWGANLQKVRILVNGKPKKVYVSARALKSGKVE 60

Query: 61 RI 62
          R+
Sbjct: 61 RV 62
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2879

A DNA sequence (GASx1869) was identified in *S.pyogenes* <SEQ ID 8257> which encodes the amino acid sequence <SEQ ID 8258>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1858(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2880

A DNA sequence (GASx1881) was identified in *S.pyogenes* <SEQ ID 8259> which encodes the amino acid sequence <SEQ ID 8260>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2752(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>

RGD motif 136-138
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF04356 GB:AF177167 type IC restriction subunit [Streptococcus thermophilus]
 Identities = 358/1047 (34%), Positives = 571/1047 (54%), Gaps = 91/1047 (8%)

Query: 7    TELELEKELIHLLEGESQWTYRKELKTEDALMDNPFKILAQNNTQYLNEEPLTASEKEQ 66
            +E  +E + I +L   E+QWTYR +LK+E+ALM NF   L + N   L E+PLT  E +Q
Sbjct: 4    SEQMIENQFIQILSEKENQWTYRPDLKSEEALMQNPRSHLNRINLAVLSEQPLTDKEFKQ 63

Query: 67   IKNQLNFVNY--YEAAKWLAGENGIAKVQVQREDAKLGTIKLEVVKADNVAGGTSVYEIA 124
            +K + + +     + A++WL GENG+A++ ++RED K   + LE  +  +++GGTS YE+
Sbjct: 64   VKVEFSRLTGTPFLASQWLRGENGVAQILLEREDGK--RVTLEAFRNKDISGGTSSYEVV 121

Query: 125  NQVAPSGSRDRRGDVTLLINGLMIQIRLKSQNHQ--CIEAFNQVKKYDKEQGPRGIPST 182
            +QV    SR  RGDV+LLINGLP+I IELK ++ +    ++A+ Q+++Y ++G F+GI++T
Sbjct: 122  HQVVPDSSRVDRGDVSLLINGLPIIHIELKKESAKDGFMQAYYQIQRYAEDGFFKGIYAT 181

Query: 183  LQMFVVSNKTDTRYIAAAKENKLNP-----NFLTQWVDQNNKFQKDLPAFAKEVLSIPRA 237
             Q+ V+SNK DTRY A   E+          FL  W  ++N+   DLF F + VL IP A
Sbjct: 182  TQIMVISNKVDTRYFARPSEDTAEAYARMKFLPNWRTEDNQTVSDLFDFTRTVLRIPDA 241

Query: 238  HQMVMTYSVIDDDKKA---LILLRPYQIHAIEAVAEASRHRKSGYIWHTTGSGKTLTSYK 294
            H+++  Y+++ DD+K    L+ LRPYQIHAI  + + +   + G+IWH TGSGKT+TS+
Sbjct: 242  HELISQYTILVDDQKNQKFLMALRPYQIHAIRKIRQKAAQHEGGFIWHATGSGKTITSFV 301

Query: 295  VARNILQIP-AVEKSIFVIDRKDLDNQTASAPGSYA---------QNDIFD--VDETEDT 342
             + + Q    V++++ V+DR DLD QT   F  +A         +N + +   +   ++
Sbjct: 302  ATKLLAQNAIGVDRTVMVVDRTDLDAQTQDEFTKFASEYHTGQTTENSVANTLIVGIKNQ 361

Query: 343  RQLIKNLESS--DRRVVVTTIQKLNAMISQMRSYDTPKFKKLKERLAHLNVVFVVDECHR 400
            +QL +NL SS  +  ++VTTIQKL+A +   +     K    E+L   ++VF+VDE HR
Sbjct: 362  KQLAQNLLSSKNNVTILVTTIQKLSAAMRSAQQESEEKGSNQFEKLRQEHTVFIVDEAHR 421

Query: 401  AVTPERQRYLTNTFRNSRWYGFTGTPIFVENNKRAQLGDLAQTTEQQYGKCLEQYTVKEAI 460
            AV+ E  + +      NS W+G TGTPIF ENK+ + G  A+TT QQYG  LH YT+K A+
Sbjct: 422  AVSDEMKRIKKLLPNSTWFGLTGTPIFPENNKQRNGTFARTTSQQYGPLLHSYTIKNAM 481

Query: 461  HDKAVLGFQVEYKTTIPD--------------MPEDS------IPEEAVDHEEHMLAVLD 500
             D AVLGFQVEY + I +                +P+D+      +P E Y+ +EH+  +L
Sbjct: 482  DDGAVLSFQVEYHSLISEEDQSVIVTQLNKGKLPDDALQQEKLLPTELYHTDEHIRTMLQ 541

Query: 501  SIINQSR--KKLGFKNGIGQTPEGLLTVKSIARAQAYYDLMKKVKAGETDLVISKKVKEK 558
             I N+    KK    NG   T   +LT  SIA+A+  Y ++K++K   T L+  ++  E+
Sbjct: 542  KIFNRRSVVKKPKVKNGF-PTMSAILLTTHSIAQAKHIYRILKEMKDNGT-LLNGRQFDER 599

Query: 559  L----PDFPKVAITYSITENDNASISRQDKMTKNLEDYNHLFGTNFTIDNLQGYNRDLND 614
                 DFP+VAIT+S   +       + D++ + +++Y   F  +   D  + YN+++N
Sbjct: 600  HQLIDKDFPKVAITFSTNPDQLEKNNEQDDELVEIMKEYEKQFDASPYQDE-KLYNQNINK 658

Query: 615  RLARKKDKFKDRHEQLDLVIVVDRLLTGFDAPCLSTIFIDRQPMKPQHIIQAFSRTNRIF 674
```

```
              RLARK+ +++   + LD VIVVDRLLTGFD+P + T++IDR+ M  Q ++QAFSRTNRI+
Sbjct: 659    RLARKEKQYQSDGQWLDFVIVVDRLLTGFDSPTIQTLYIDRE-MNYQKLLQAFSRTNRIY 717

Query: 675   ESRGHYGQVVTTQTPLRFKEAVDKALSLYSNGGEN-DVLAP-EWEERKARPFEKVTVLKN 732
               + K  G +V+F+ P    +E V      L+SN  +N D L P  +EE K  F E  T+ K
Sbjct: 718    -TGKDGGLIVSFRKPFTMRENVRNTPRLPSNKNQNFDQLIPKEYEEVKKEFIECSTLYKQ 776

Query: 733   IVPDPDAPFTIESAQTAFLKQYAKAFQAPDKLPASVQVVYSDFNETLLSEVGLSDEVIDTY 792
                 D    P       A +  Y K  +++  L +  Q   DF E   SEV    E +  Y
Sbjct: 777    SSADLSDNPNDLKTMIAQVSAYQKLEKSYKALRSYDQYEHDFEE--FSEV---VEQLPQY 831

Query: 793   KGTVQNVIAEITRKRRED--------DEAIPEINIDYELESVQMDDINYHYILTLIQAFVD 844
               +G  +N+  +I++  ED          ++ + EI     L +   D ++  YI  L++A
Sbjct: 832    QGKTENIKTKIKSMIEDEGHPEDPEKLLQEIAFSSQLNATHKDVVDSFYINQLLKAIQL 891

Query: 845   QEQEALQERLNDNMDQYIQDLAKSNPAMADSLAELWQDIQEPKAYEGKSIVYELDNLI 904
               E  A+++   + +Q    + K    + D L    ++I          +    +  I
Sbjct: 892    NEAGAVEK--FRKEIQQKDPQIQKMYHTLKDQLVNTTEEI----------DVAQLKETSI 939

Query: 905   GDKIQRAIKHFADQWKADPDKLAFVATNYHSANSTKQVGMSTLKE-SLDYQAYKEKQGDS 963
               ++IQR ++  A+++   D L      Y S  T       L  +L  + ++ K G+
Sbjct: 940    QNEIQRQLQKEAEEPGLSFDFLQSAMNEYQSDKETIPYLTHLLDSMTLSKEEPEARTGE- 998

Query: 964   AMNKLKYKSQFERELVQFIRDQIQPLK 990
                 K + +++    E +Q   +Q+Q  K
Sbjct: 999    ---KYRRRTKVLEERLQQNFEQLQKWK 1022
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2881

A DNA sequence (GASx1882) was identified in *S.pyogenes* <SEQ ID 8261> which encodes the amino acid sequence <SEQ ID 8262>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.3653(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB53491 GB:U35629 unknown [Lactococcus lactis subsp. lactis]
 Identities = 141/241 (58%), Positives = 178/241 (73%)

Query: 3    KSKQPQYRVDGIFKGWWEKRLGDIVQITMGQSPSSQNVTTNPSDYILVQGNADIKNGYVF 62
             K K P+ RP GF  EWE ++LGD V+I MGQSP+S+NYT +P+DYILVQGNAD+KNG V
Sbjct: 13   KKKVPELRFKGFTDEWELRKLGDEVKIVMGQSPNSNYTYDPNDYILVQGNADMKNGRVL 72

Query: 63   PRVWTTQITKQADKGDILLSVRAPVGDVGKTNYHVIIGRGVAAIKGNEFIPQYLKYLKEI 122
             PRVWTTQ+TKQA+K D+ILSVRAPVGD+GKT Y V+IGRGVAAIKGNEFIPQ L  +K
Sbjct: 73   PRVWTTQVTKQARKDDLILSVRAPVGDIGKTAYDVVIGRGVAAIKGNEFIPQNLGKMKSD 132

Query: 123  GYWKRISTGSTTFDSISSSDIKYAKIQIPSLPEQRAIGRLFQMVDQLIQLQDQKLATLKEQ 182
             GYW R STGSTTF+SI+S+DIK A I +P++ EQ+ IG  F+ +D  I L  +KL  LKEQ
Sbjct: 133  GYWTRYSTGSTTFESINSTDIKEAIISVPAIEEQDKIGSFFKQLDNTIALHQKKLDLLKEQ 192

Query: 183  KQTFLRKMFPAQGQKVPEIRLQGFKGEWEEKKLREVSTHRSGTAIEKYFDSEGEFKVISIG 243
             K+ FL+KMFP  G KVPE+R  GF  +WEE+KL +++   +G          G++  + G
Sbjct: 193  KNGFLQKMFPKNGAKVPELRFAGFADDWEERKLGDITKISTGKLDANAMVENGKYDFYTSG 253
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2882

A DNA sequence (GASx1883) was identified in *S.pyogenes* <SEQ ID 8263> which encodes the amino acid sequence <SEQ ID 8264>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4318(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF94357 GB:AF177167 type IC modification subunit [Streptococcus thermophilus]
 Identities = 293/523 (56%), Positives = 377/523 (72%), Gaps = 6/523 (1%)

Query: 6   TSLRQALWHSADQLRGQMDANDYKNYLLGLIFYKHLSDKLLIAVCDNLEKHFNTFTEAQK 65
           T+L Q LW SAD LRG+MDA++YKNYLLGLIFYK+LSDK L  V +       +TF E
Sbjct: 3   TSLNQQLWASADILRGKMDASEYKNYLLGLIFYKYLSDKQLREVYEQENGKTDTFPERST 62

Query: 66  I---FEDAYQDESLKDDLISVVTGDLGYFIEPTLTFEKLIQDVYHNTPQLESLAQGFRDI 122
           +   F + Y+++  KDDLI  +    GYFI+P   F        +  F L  L  GF ++
Sbjct: 63  LYAGFMEWYSED--KDDLIENIQPRQGYFIQPDRLFYHYRIKADNYEFNLTDLQAGFNEL 120

Query: 123 EQSGEDFENLFEDIDLYSKKLGSTPQKQNQTISNVMKTLNEIDFEAVDGYTLGDAYEYLI 182
           E+ GE+F  LF DIDL S KLGS  Q++N TI+ V++ L+EID     +GD +GDAYEYLI
Sbjct: 121 ERQGESFSGLFSDIDLNSTKLGSNAQQRNVTITEVLRALDEIDLFEHNGDVIGDAYSYLI 180

Query: 183 GEFASESGKKAGEFYTPQAVSHLMTQIVFLGREDQKGMTLYDPAMGSGSLLLNAKKYSNQ 242
           G FA+ +GKKAGEFYTPQAVS +M++I  +G+E +     +YDPAMGSGSL+LN ++Y
Sbjct: 181 GMFAAGAGKKAGEFYTPQAVSRIMSEITSIGQESRVPFHIYDPAMGSGSLMLNIRRYLIH 240

Query: 243 SDTVSYYGQEINTSTYNLARMNMMLHGVAIENQHLSNADTLDADWPTDEPINFDGVLMNP 302
            + V Y+GQE+NT+T+NLARMN++LHGV  E  +L+N DTLDADWP++EP  FD V+MNP
Sbjct: 241 PNQVHYHGQELNTTTFNLARMNLILHGVDKERMNLNNGDTLDADWPSEEPYQFDSVVMNP 300

Query: 303 PYSLKWSATAGFLTDPRFSSYGVLAPKSKADFAFLLHGFYHLKNTGTMAIVLPHGVLFRG 362
           PYS KWSA   FL+DPRF  +G LAPKSKADFAFLLHGFYHLK +GTM IVLPHGVLFRG
Sbjct: 301 PYSAKWSAADKFLSDPRERFGKLAPKSKADFAFLLHGFYHLKESGTMGIVLPHGVLFRG 360

Query: 363 AAEGKIRQKLLEQGAIDTIIGLPSNIFYNTSIPTTIILLKKNRTNKDVFFIDASKEFDKG 422
            AEG IRQ LLE GAID +IGLP+NIF+ TSIPTT+IILKKN+ +DV FIDAS++F+K
Sbjct: 361 GAEGTIRQALIEMGAIDAVIGLPANIFGTSIPTTVIILKKNRSRRDVLFIDASQDFEKQ 420

Query: 423 KNQNTMTDNHIKKILDAYKSRDNSDKPSYLASFDEIIENDYNLNIPRYVDTFEEVPVKPL 482
           KNQN + D HI KI+  YK R++ ++++++ASFDEI END+NLNIPRYVDTFEE      L
Sbjct: 421 KNQNVLLDEHIDKIVSTYKKRRDIERYAHVASFDEIQENDFNLNIPRYVDTFEEEEPVDL 480

Query: 483 PELAKQLSDIDQKIAKTNAKLDQLMKQLVGTTKEAQDELDTPR 525
            E+   L  I++E+ +    L  L+     ++E Q  +++ R
Sbjct: 481 VEVTNLLKINSELVQQEQTLLSLINDF-SESSENQAMIESMR 522
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2883

A DNA sequence (GASx1886R) was identified in *S.pyogenes* <SEQ ID 8265> which encodes the amino acid sequence <SEQ ID 8266>. Analysis of this protein sequence reveals the following:

```
Possible site: 59

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -8.17   Transmembrane  155 - 171 ( 147 - 173)
   INTEGRAL    Likelihood = -7.22   Transmembrane   14 -  30 (  11 -  33)
   INTEGRAL    Likelihood = -7.17   Transmembrane  182 - 198 ( 179 - 205)
   INTEGRAL    Likelihood = -5.68   Transmembrane  132 - 148 ( 128 - 152)
   INTEGRAL    Likelihood = -4.14   Transmembrane   46 -  62 (  43 -  62)
   INTEGRAL    Likelihood = -3.50   Transmembrane   73 -  89 (  73 -  90)
   INTEGRAL    Likelihood = -0.96   Transmembrane   95 - 111 (  95 - 111)

----- Final Results -----
              bacterial membrane --- Certainty=0.4270(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2884

A DNA sequence (GASx1890R) was identified in *S.pyogenes* <SEQ ID 8267> which encodes the amino acid sequence <SEQ ID 8268>. Analysis of this protein sequence reveals the following:

```
Possible site: 58

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4757(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>

RGD motif 339-341
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA62650 GB:L37110 clyM [Plasmid pAD1]
 Identities = 127/492 (25%), Positives = 230/492 (45%), Gaps = 30/492 (6%)

Query: 46  KLFYSEPENQLPETIMFLSMKTLVLDINHPSKRIENK----SRAYEQYIQQ-IREENGIN 100
           K F       L + ++ L+ KTLVLD++ F K    K    S+ +  Y+++    +  I
Sbjct: 135 KEFIINLLENLTQELIHLTSKTLVLDLHTPKKNEPLKGNDSSKRFIYYLKKRFNSKKDII 194

Query: 101 HFFDRYPYLLKQINKEVGLIEESYSLLFDRFLEDLSEIKSCFNI-SEPLSNVAFSLGDSH 159
             F+  YP L++     +    ++   +  R  EDL  I++CFNI S  L+++++ S GDSH
Sbjct: 195 AFYTCYPELMRITVVRMKYFLDNFKQMLIRVTEDLPSIQNCFNIQSSELNSISESQSDSH 254

Query: 160 SKKQTVVKIAFKE-KSVYYKPKSYHSHSILLKLTSLLKSSNIPSFSLPKSLVKADYCWQL 218
           S+ +TV  + F + K + YKPK  +S + L +      L       + K + +  Y ++
Sbjct: 255 SRGKTVSTLTFSDGKKIVYKPK-INSENKLRDFFEFLNKKLEADIYIVKKVIRNTFYYES 313

Query: 219 GVAYTSSNK-DEVAKIYFKYGVLAAFSKIPSITTDLIMENVIVSGGDLYLLIDVETFPQRKL 277
            +     N  +EV K Y +YG L   + +F++TDLH EN+I  G    +ID ETFPQ+ +
Sbjct: 314 YIDNIEINNIEKVKKYYERYGKLIGIAFLFNVTDLHYENIIAHGEYPVIIDNETFFQQNI 373
```

```
Query: 278 NVQNQNPEGITVDTYQRIYETSLSNGLFF---VQFEKNSAPNVSGISRKGGKRQKGKYEL 334
           ++  N    TVD   +  ++ +  GL F   ++  + +S      +S     K Q    +++
Sbjct: 374 PIEFGN--SATVDAKYKYLDSIMVTGLVPYLAMKDKSDSKDEGVNLSALNFKEQSVPFKI 431

Query: 335 I---NKNRGDLKLVKVDYFQKDRFNIPTLNGKVVKPLDYANEIISGFRECYIFLLSQRSK 391
           +   N    +++      +  +   N P +N + +  + Y    I++G +    +      + K
Sbjct: 432 LKIKNTFTDEMRPEYQTHIMDTAKNTPIMNNEKISPISYEKYIVTGMKSILMKAKDSKKK 491

Query: 392 IKEIV-EGPPELKSRVDPRNTSDYGKPLQASTNPKYLFS----EKKRKNLFSILYETKHI 446
           I   +      L  R   R T  Y   L+ S +P     +    EK   N+++  Y+ K +
Sbjct: 492 ILAYINNNLQNLIVRNVIRPTQKYADMLEFSYHPNCFSNAIKREKVLHNMWAYPYKNKKV 551

Query: 447 EHFIVDNEIKDLMNGDIP-YFSMDTRGNVYNSVGTLIGNLGDTTSL---PDSITILNDER 502
            H+     E  DL++GDIP +++   ++ ++  S G L+ +     ++L     + I   L DE
Sbjct: 552 VHY----EFSDLIDGDIPIFYNNISKTSLIASDGCLVKDFYQESALNRCLNKINDLCDED 607

Query: 503 LKFTCELLEIVL 514
           +      LE  L
Sbjct: 608 ISIQTVWLEILAL 619
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2885

A DNA sequence (GASx1891R) was identified in *S.pyogenes* <SEQ ID 8269> which encodes the amino acid sequence <SEQ ID 8270>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.3487(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA02867 GB:L07740 salivaricin A [Streptococcus salivarius]
 Identities = 46/51 (90%), Positives = 48/51 (93%)

Query: 1  MSFMKNSKDILTNAIEEVSEKELMEVAGGKKGSGWFATITDDCPNSVFVCC 51
          M+ MKNSKDIL NAIEEVSEKELMEVAGGK+GSGW ATITDDCPNSVFVCC
Sbjct: 1  MNAMKNSKDILNNAIEEVSEKELMEVAGGKRGSGWLATITDDCPNSVFVCC 51
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2886

A DNA sequence (GASx1901R) was identified in *S.pyogenes* <SEQ ID 8271> which encodes the amino acid sequence <SEQ ID 8272>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.59   Transmembrane    3 -  19 (   1 -  20)

----- Final Results -----
               bacterial membrane --- Certainty=0.1638(Affirmative) < succ>
```

```
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2887

A DNA sequence (GASx1905R) was identified in *S.pyogenes* <SEQ ID 8273> which encodes the amino acid sequence <SEQ ID 8274>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -0.48   Transmembrane   38 -  54 (  37 -  54)

----- Final Results -----
               bacterial membrane --- Certainty=0.1192(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2888

A DNA sequence (GASx1911R) was identified in *S.pyogenes* <SEQ ID 8275> which encodes the amino acid sequence <SEQ ID 8276>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-10.40   Transmembrane   27 -  43 (  22 -  48)
   INTEGRAL    Likelihood = -9.82   Transmembrane   52 -  68 (  50 -  74)
   INTEGRAL    Likelihood = -7.27   Transmembrane  113 - 129 ( 111 - 134)
   INTEGRAL    Likelihood = -1.97   Transmembrane  137 - 153 ( 135 - 153)

----- Final Results -----
               bacterial membrane --- Certainty=0.5161(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2889

A DNA sequence (GASx1915R) was identified in *S.pyogenes* <SEQ ID 8277> which encodes the amino acid sequence <SEQ ID 8278>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-10.77   Transmembrane  242 - 258 ( 238 - 262)

----- Final Results -----
                bacterial membrane --- Certainty=0.5310(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2890

A DNA sequence (GASx1918R) was identified in *S.pyogenes* <SEQ ID 8279> which encodes the amino acid sequence <SEQ ID 8280>. Analysis of this protein sequence reveals the following:

```
Possible site: 38

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -7.32   Transmembrane   40 -  56 (  39 -  60)

----- Final Results -----
                bacterial membrane --- Certainty=0.3930(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2891

A DNA sequence (GASx1923R) was identified in *S.pyogenes* <SEQ ID 8281> which encodes the amino acid sequence <SEQ ID 8282>. Analysis of this protein sequence reveals the following:

```
Possible site: 42

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood =-12.26   Transmembrane   20 -  36 (  13 -  42)

----- Final Results -----
                bacterial membrane --- Certainty=0.5904(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2892

A DNA sequence (GASx1926) was identified in *S.pyogenes* <SEQ ID 8283> which encodes the amino acid sequence <SEQ ID 8284>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2322(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2893

A DNA sequence (GASx1928R) was identified in *S.pyogenes* <SEQ ID 8285> which encodes the amino acid sequence <SEQ ID 8286>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3395(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2894

A DNA sequence (GASx1929R) was identified in *S.pyogenes* <SEQ ID 8287> which encodes the amino acid sequence <SEQ ID 8288>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -1.86   Transmembrane   17 -  33 (  15 -  33)

----- Final Results -----
             bacterial membrane --- Certainty=0.1744(Affirmative) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2895

A DNA sequence (GASx1931R) was identified in *S.pyogenes* <SEQ ID 8289> which encodes the amino acid sequence <SEQ ID 8290>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0551(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2896

A DNA sequence (GASx1941R) was identified in *S.pyogenes* <SEQ ID 8291> which encodes the amino acid sequence <SEQ ID 8292>. Analysis of this protein sequence reveals the following:

```
Possible site: 16

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2377(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2897

A DNA sequence (GASx1949) was identified in *S.pyogenes* <SEQ ID 8293> which encodes the amino acid sequence <SEQ ID 8294>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0262(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2898

A DNA sequence (GASx1951R) was identified in *S.pyogenes* <SEQ ID 8295> which encodes the amino acid sequence <SEQ ID 8296>. Analysis of this protein sequence reveals the following:

```
Possible site: 45

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1330(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2899

A DNA sequence (GASx1953) was identified in *S.pyogenes* <SEQ ID 8297> which encodes the amino acid sequence <SEQ ID 8298>. Analysis of this protein sequence reveals the following:

```
Possible site: 15

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2900

A DNA sequence (GASx1957) was identified in *S.pyogenes* <SEQ ID 8299> which encodes the amino acid sequence <SEQ ID 8300>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2409(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2901

A DNA sequence (GASx1969) was identified in *S.pyogenes* <SEQ ID 8301> which encodes the amino acid sequence <SEQ ID 8302>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL     Likelihood = -2.28   Transmembrane     7 -  23 (   7 -  23)

----- Final Results -----
               bacterial membrane --- Certainty=0.1914(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2902

A DNA sequence (GASx1971R) was identified in *S.pyogenes* <SEQ ID 8303> which encodes the amino acid sequence <SEQ ID 8304>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1545(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2903

A DNA sequence (GASx1973) was identified in *S.pyogenes* <SEQ ID 8305> which encodes the amino acid sequence <SEQ ID 8306>. Analysis of this protein sequence reveals the following:

```
Possible site: 49

>>> Seems to have no N-terminal signal sequence
   INTEGRAL     Likelihood = -2.44   Transmembrane    31 -  47 (  31 -  48)

----- Final Results -----
               bacterial membrane --- Certainty=0.1977(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB51744 GB:AJ245405 speX [Streptococcus pyogenes]
 Identities = 236/256 (92%), Positives = 243/256 (94%)

Query: 3   MIISFESVILKHNKIITPEKRLPMKKTKLIFSFTSIFIAIISRPVFGLEVDNNSLLRNIY 62
           MIISFESVILKHNKIITPEKRLPMKKTKLIFSFTSIFIAIISRPVFGLEVDNNSLLRNIY
Sbjct: 1   MIISFESVILKHNKIITPEKRLPMKKTKLIFSFTSIFIAIISRPVFGLEVDNNSLLRNIY 60

Query: 63  STIVYEYSDTVIDFKTSHNLVTKKLDVRDARDFFINSEMDEYAANDFKDGDKIAMFSVPF 122
           STIVYEYSD VIDFKTSHNLVTKKLDVRDARDFFINSEMDEYAANDFK GDKIA+FSVPF
Sbjct: 61  STIVYEYSDIVIDFKTSHNLVTKKLDVRDARDFFINSEMDEYAANDFKTGDKIAVFSVPF 120

Query: 123 DWNYLSEGKVIAYTYGGMTPYQEEPMSKNIPVNLWINRKQIPVPYNQISTNKTTVTAQEI 182
           DWNYLS+GKV AYTGG+TPYQ+    K   VNLWIN KQI VPYN+ISTNKTTVTAQEI
Sbjct: 121 DWNYLSKGKVTAYTGGITPYQKLQYLKISLVNLWINGKQISVPYNEISTNKTTVTAQEI 180

Query: 183 DLKVRKFLISQHQLYSSGSSYKSGKLVFHTNDNSDKYSLDLFYVGYRDKESIFKVYKDNK 242
           DLKVRKFLI+QHQLYSSGSSYKSG+LVFHTNDNSDKYS DLFYVGYRDKESIFKVYKDNK
Sbjct: 181 DLKVRKFLIAQHQLYSSGSSYKSGRLVFHTNDNSDKYSFDLFYVGYRDKESIFKVYKDNK 240

Query: 243 SFNIDKIGHLDIEIDS 258
           SFNIDKIGHLDIEIDS
Sbjct: 241 SFNIDKIGHLDIEIDS 256
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2904

A DNA sequence (GASx1974R) was identified in *S.pyogenes* <SEQ ID 8307> which encodes the amino acid sequence <SEQ ID 8308>. Analysis of this protein sequence reveals the following:

```
Possible site: 53

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2022(Affixmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2905

A DNA sequence (GASx1983) was identified in *S.pyogenes* <SEQ ID 8309> which encodes the amino acid sequence <SEQ ID 8310>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
```

```
bacterial cytoplasm --- Certainty=0.0989(Affirmative) < succ>
 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2906

A DNA sequence (GASx1987) was identified in *S.pyogenes* <SEQ ID 8311> which encodes the amino acid sequence <SEQ ID 8312>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2389(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2907

A DNA sequence (GASx1988) was identified in *S.pyogenes* <SEQ ID 8313> which encodes the amino acid sequence <SEQ ID 8314>. Analysis of this protein sequence reveals the following:

```
Possible site: 48

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.5904(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB16031 GB:AB030747 transposase [Streptococcus pyogenes]
 Identities = 22/24 (91%), Positives = 23/24 (95%)

Query: 1   LERLPGTAKEYHNLCYTREKGKSK 24
           +ERLPGTAKEYHNL YTREKGKSK
Sbjct: 399 IERLPGTAKEYHNLRYTREKGKSK 422
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2908

A DNA sequence (GASx1990R) was identified in *S.pyogenes* <SEQ ID 8315> which encodes the amino acid sequence <SEQ ID 8316>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
               bacterial outside --- Certainty=0.3000(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2909

A DNA sequence (GASx1991) was identified in *S.pyogenes* <SEQ ID 8317> which encodes the amino acid sequence <SEQ ID 8318>. Analysis of this protein sequence reveals the following:

```
Possible site: 53

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -0.16   Transmembrane    2 -  18 (   1 -  18)

----- Final Results -----
               bacterial membrane --- Certainty=0.1065(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2910

A DNA sequence (GASx1994) was identified in *S.pyogenes* <SEQ ID 8319> which encodes the amino acid sequence <SEQ ID 8320>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -1.44   Transmembrane   28 -  44 (  28 -  44)

----- Final Results -----
               bacterial membrane --- Certainty=0.1574(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2911

A DNA sequence (GASx1996) was identified in *S.pyogenes* <SEQ ID 8321> which encodes the amino acid sequence <SEQ ID 8322>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1076(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2912

A DNA sequence (GASx1997R) was identified in *S.pyogenes* <SEQ ID 8323> which encodes the amino acid sequence <SEQ ID 8324>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -7.96   Transmembrane   53 -  69 (  49 -  75)
   INTEGRAL    Likelihood = -2.34   Transmembrane   24 -  40 (  24 -  43)

----- Final Results -----
               bacterial membrane --- Certainty=0.4185(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2913

A DNA sequence (GASx2007R) was identified in *S.pyogenes* <SEQ ID 8325> which encodes the amino acid sequence <SEQ ID 8326>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -6.64   Transmembrane   46 -  62 (  43 -  65)

----- Final Results -----
               bacterial membrane --- Certainty=0.3654(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB97959 GB:U96166 ATP-binding cassette lipoprotein
          [Streptococcus cristatus]
 Identities = 37/60 (61%), Positives = 42/60 (69%), Gaps = 1/60 (1%)

Query: 59  FLTACGTKKDSKKKEEVKEIKMSDIKDDAVSKTKVVDGEEVIEYTTKDGNVIQIPAGNEE 118
           FL ACG+K     KE + + K  D K DAV +KTK VDG+EVIEYT  DGNVIQIPA  EE
Sbjct: 12  FLAACGSKNADNKE-ISDGKKVDFKKDAVDQKTKTVDGKEVIEYTMPDGNVIQIPADGEE 70
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2914

A DNA sequence (GASx2009) was identified in *S.pyogenes* <SEQ ID 8327> which encodes the amino acid sequence <SEQ ID 8328>. Analysis of this protein sequence reveals the following:

```
Possible site: 41

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1246(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2915

A DNA sequence (GASx2010) was identified in *S.pyogenes* <SEQ ID 8329> which encodes the amino acid sequence <SEQ ID 8330>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2549(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2916

A DNA sequence (GASx2012R) was identified in *S.pyogenes* <SEQ ID 8331> which encodes the amino acid sequence <SEQ ID 8332>. Analysis of this protein sequence reveals the following:

```
Possible site: 28
```

```
>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.3307(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA27007 GB:L26141 pyrogenic exotoxin B [Streptococcus pyogenes]
 Identities = 40/102 (39%), Positives = 57/102 (55%), Gaps = 7/102 (6%)

Query: 2   EMHFVRTEPEARRIAETFCAENTQTKTPMRVQQLSYPSDTDHSGGEL-----YIYALSPA 56
           + +F R E EA+  A TF  ++    K   R  +     D  + GGEL     YIY +S
Sbjct: 28  DQNFARNEKEAKDSAITFIQKSAAIKNGARSAE-DIKLDKVNLGGELSGSNMYIYNISTG 86

Query: 57  GFIIVSGDTRAHTILGYSFDNNLDIN-HDNVRSMIEAYQKQI 97
           GF+IVSGD R+  ILGYS   + D+N  +N+ S +E+Y +QI
Sbjct: 87  GFVIVSGDKRSPEILGYSTGSGFDVNCKENIASPMESYVRQI 128
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2917

A DNA sequence (GASx2013R) was identified in *S.pyogenes* <SEQ ID 8333> which encodes the amino acid sequence <SEQ ID 8334>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
                bacterial outside --- Certainty=0.3000(Affirmative) < succ>
                 bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2918

A DNA sequence (GASx2014R) was identified in *S.pyogenes* <SEQ ID 8335> which encodes the amino acid sequence <SEQ ID 8336>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.1392(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2919

A DNA sequence (GASx2015) was identified in *S.pyogenes* <SEQ ID 8337> which encodes the amino acid sequence <SEQ ID 8338>. Analysis of this protein sequence reveals the following:

```
Possible site: 35

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL     Likelihood = -1.75   Transmembrane   18 -  34 (  17 -  37)

----- Final Results -----
                bacterial membrane --- Certainty=0.1702(Affirmative) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2920

A DNA sequence (GASx2018) was identified in *S.pyogenes* <SEQ ID 8339> which encodes the amino acid sequence <SEQ ID 8340>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence
   INTEGRAL     Likelihood = -5.84   Transmembrane   23 -  39 (  22 -  40)

----- Final Results -----
                bacterial membrane --- Certainty=0.3336(Affirmative) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2921

A DNA sequence (GASx2019) was identified in *S.pyogenes* <SEQ ID 8341> which encodes the amino acid sequence <SEQ ID 8342>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.0669(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                  bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC98898 GB:AF023179 low temperature requirement C protein
           [Listeria monocytogenes]
 Identities = 95/144 (65%), Positives = 117/144 (80%)

Query: 15  LAERGVSLEAIAELVLFLQNDYIPNLTMARCLESVEAVLAKREVQNAIITGVELDKLAEA 74
           L ERGV ++ IAELVLFLQ  Y P L +  C ++VE VL KREVQNA++TG++LD +AE
Sbjct: 16  LIERGVEIDDIAELVLFLQQKYHPGLELDICRQNVEHVLRKREVQNAVLTGIQLDVMAEK 75

Query: 75  NQLSEPLLSILKTDQGLYGIDEILALSIVNLYGSIGFTNYGYLDKTKPGIVDKLNHKDGY 134
            +L +PL +I+  D+GLYG+DEILALSIVN+YGSIGFTNYGY+DK KPGI+ KLN  DG
Sbjct: 76  GELVQPLQNIISADEGLYGVDEILALSIVNVYGSIGFTNYGYIDKVKPGILAKLNEHDGI 135

Query: 135 SCHTFLDDIVSAIAAAAASRIAHN 158
           + HTFLDDIV AIAAAAASR+AH+
Sbjct: 136 AVHTFLDDIVGAIAAAAASRLAHS 159
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2922

A DNA sequence (GASx2030) was identified in *S.pyogenes* <SEQ ID 8343> which encodes the amino acid sequence <SEQ ID 8344>. Analysis of this protein sequence reveals the following:

```
Possible site: 18

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.0320(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2923

A DNA sequence (GASx2031) was identified in *S.pyogenes* <SEQ ID 8345> which encodes the amino acid sequence <SEQ ID 8346>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.0583(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2924

A DNA sequence (GASx2032R) was identified in *S.pyogenes* <SEQ ID 8347> which encodes the amino acid sequence <SEQ ID 8348>. Analysis of this protein sequence reveals the following:

```
Possible site: 53

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -2.76   Transmembrane   27 -  43 (  26 -  43)

----- Final Results -----
               bacterial membrane --- Certainty=0.2105(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

A related GBS gene <SEQ ID 8467> and protein <SEQ ID 8468> were also identified. Analysis of this protein sequence reveals the following:

```
Lipop: Possible site: -1   Crend: 10
McG: Discrim Score:    -11.19
GvH: Signal Score (-7.5): -4.94
     Possible site: 49
>>> Seems to have no N-terminal signal sequence
ALOM program   count: 1 value:  -4.19 threshold:  0.0
   INTEGRAL    Likelihood = -4.19   Transmembrane   25 -  41 (  25 -  42)
   PERIPHERAL  Likelihood = 13.26      41
 modified ALOM score:   1.34

*** Reasoning Step: 3

----- Final Results -----
               bacterial membrane --- Certainty=0.2678(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

The protein has homology with the following sequences in the databases:

```
ORF01616(304 - 429 of 771)
SP|O06442|SECE_STAAU(7    -   48    of    60)    PREPROTEIN    TRANSLOCASE    SECE    SUBUNIT.
GP|2078376|gb|AAB54017.1||U96619 SecE {Staphylococcus aureus}
%Match = 5.4
%Identity = 26.2  %Similarity = 57.1
Matches = 11  Mismatches = 18  Conservative Sub.s = 13

99        129        159        189        219        249        279        309
RIIQIMLK*HLWRRYGTKESKPSVYRMRKPKLLNRSK*HPQANTTRSK*IL*IL*EVYNTQRNALI*RNKLQKGELIMPV
                                                                              |
                                                                              MAKKESFF

339        369        399        429        459        489        519        549
KGIFQVLRDTTWPNKKQRWKDFISILEYTVPFTIVIYIPDKLAAGVMDLINRF***IILDRNNPNP*ILLRVFCVENNI
||:    :  |:|| ::: :|  : ::   :|| :  | :|
KGVKSEMEKTSWPTKEELFKYTVIVVSTVIPFLVFFYALDLGITALKNLLFG
          20         30         40         50         60
```

SEQ ID 8468 (GBS396) was expressed in *E.coli* as a GST-fusion product. SDS-PAGE analysis of total cell extract is shown in Figure 83 (lane 9; MW 35kDa).

GBS396-GST was purified as shown in Figure 217, lane 8.

Example 2925

A DNA sequence (GASx2034R) was identified in *S.pyogenes* <SEQ ID 8349> which encodes the amino acid sequence <SEQ ID 8350>. Analysis of the protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.59   Transmembrane   53 -  69 (  53 -  70)

----- Final Results -----
               bacterial membrane --- Certainty=0.1235(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2926

A DNA sequence (GASx2035) was identified in *S.pyogenes* <SEQ ID 8351> which encodes the amino acid sequence <SEQ ID 8352>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2928(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2927

A DNA sequence (GASx2042R) was identified in *S.pyogenes* <SEQ ID 8353> which encodes the amino acid sequence <SEQ ID 8354>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2547(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2928

A DNA sequence (GASx2043) was identified in *S.pyogenes* <SEQ ID 8355> which encodes the amino acid sequence <SEQ ID 8356>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3289(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2929

A DNA sequence (GASx2049) was identified in *S.pyogenes* <SEQ ID 8357> which encodes the amino acid sequence <SEQ ID 8358>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4014(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2930

A DNA sequence (GASx2052) was identified in *S.pyogenes* <SEQ ID 8359> which encodes the amino acid sequence <SEQ ID 8360>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
```

```
bacterial outside --- Certainty=0.3000(Affirmative) < succ>
bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2931

A DNA sequence (GASx2055R) was identified in *S.pyogenes* <SEQ ID 8361> which encodes the amino acid sequence <SEQ ID 8362>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3048(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:BAB05703 GB:AP001513 imidazolonepropionase
          (imidazolone-5-propionate hydrolase) [Bacillus halodurans]
 Identities = 203/416 (48%), Positives = 278/416 (66%), Gaps = 4/416 (0%)

Query: 11  DVLLTHFNQLPCLNDPGHPLTGQEMKKATIVEDGYIAIKDGLIVALGSGEPDAELVGTQT 70
           D LL +  QL  +   G P  G+EM +  ++E   + I+DG +  +G+          Q
Sbjct: 6   DTLLVNIGQLLPMESKG-PKRGKEMSELQLLEHAALGIRDGKVAFIGTMVEADTFTANQM 64

Query: 71  IMRSYKGKIATPGIIDCHTHLVYGGSREHEFAKKLAGVSYLDILAQGGGILSTVRATRSA 130
           I    +GK+ TPG++D HTHL++GGSREHE A K  GV YL+IL  GGGILETV ATR+A
Sbjct: 65  I--DCQGKLVTPGLVDPHTHLIFGGSREHEMALKQQGVPYLEILKNGGGILATVEATRAA 122

Query: 131 SFDNLYQKSKRLLDYMLLHGVTTVEAKSGYGLDWETEKRQLDVVAALEKDHPIDLVSTFM 190
           S + L  K+   L+ ML +GVTT+EAKSGYGLD ETE +QL    A+ + HPID+VSTF+
Sbjct: 123 SEEELITKAICHLNRMLSYGVTTIEAKSGYGLDRETEWKQLRAAKAVGEQHPIDIVSTFL 182

Query: 191 AAHAIPEEYKGNPKAYLDVIIKDMLPVVKEENLAEFCDIFCEKNVFTADESRYLLSKAKE 250
            AHAIP  ++ +P  +LD +  DML  +KE+NLAEF DIF E  VFT ++SR  L KAKE
Sbjct: 183 GAHAIPTSHRNDPDRFLDEMA-DMLGEIKRQNLAEFVDIFTETGVFTVEQSRTFLQKAKE 241

Query: 251 MGFKLRIHADEIASIGGVDVAAELGAVSADHLMMITDDGIAKLIGAGVIGNLLPATTFSL 310
            GF L+++HADEI  +GG ++A KL A+SA+HL+  +D GI K+  AG I  LLP TTF L
Sbjct: 242 RGFGLRLHADEIDPLGGAELAGELGAISADHLVGASDQGIQKMAAAGTIACLLPGTTFYL 301

Query: 311 MEDTYAPARKMIDAGMATTLSTDSNPGSCPTANMQFVMQLGCFMLRLFTPIEVLNAVTINA 370
           +DTYA AR MID G+A+T+STD NPGS PT N+Q +M +     L++TP E+ +AVT+N
Sbjct: 302 GKDTYAARARMIDQGLAVTISTDFNPGSSPTENLQLIMSIAARLKMTPEEIWHAVTVNG 361

Query: 371 AYSVNRQERVGSLTVGKRADIAIFDAPNIDYPFYFPATNLIHQVYKKGQLTVDRGR 426
           A+++ R +  G L VG+ AD+ ++DA N  Y  Y +  N +H V+KKG++  +R R
Sbjct: 362 AHAIGRGDTAGQLAVGRAADVVVWDAKNYYYVPVHHGVNHVHSVWKKGEVVYERRR 417
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2932

A DNA sequence (GASx2056) was identified in *S.pyogenes* <SEQ ID 8363> which encodes the amino acid sequence <SEQ ID 8364>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1847(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB61139 GB:AL132952 predicted using Genefinder-cDNA EST
           yk155e6.3 comes from this gene-cDNA EST yk155e6.5 comes
           from this gene-cDNA EST yk156d6.5 comes from this
           gene-cDNA EST yk259b10.3 comes fr
 Identities = 302/649 (46%), Positives = 419/649 (64%), Gaps = 17/649 (2%)

Query: 29  EGIRRAPDRGPRLTQAQTEIALKNALRYVPTKFHSEVIPEFLEELKTRGRIYGYRFRPKD 88
           + +  AP R   LTQ +  +A++NALRY+P + H  +  EF EKL T G IYGYRF P
Sbjct: 85  KNVAHAPKRPCNLTQTEKMLAVRNALRYIPKEHHVLLATEPAEELNTYGHIYGYRPMPNF 144

Query: 89  RIYGKPIDEYKGNCTAAKAMQVMIDNNLSFELALYPYELVTYGETGSVCANWMQYCLIKK 148
            ++  P+ E   +C  A A+ +MI NNL    +A +P ELVTYG  G V +NW+Q+ L+ +
Sbjct: 145 DLFAPPVSEIGAHCBQASAIILMILNNVLDKRVACPFQELVTYGGNSQVFSNWIQPRLVLR 204

Query: 149 YLEVMTDBQTLVVESGHPVGLPKSKPEAPRVIITNGLLVGEYDNMKDWEIAEEMGVTNYG 208
           YL  M1D QTLV+ SGHP+GLP S P++PR+ +TNG+++  Y   + ++     +GVT YG
Sbjct: 205 YLVTMTDHQTLVLYSGHPLGLPPSTPDSPRMTVTNGMMIPSYSTKELVDKYFALGVTQYG 264

Query: 209 QMTAGGWMYIGPQGIVHGTFNTLLNAGRLKLGVADDGDLTGKLPISSGLGGMSGAQGKAA 268
           QMTAG + YIGPQGIVHGT  T+LNAGR ++G+     L GK+F+++GLGGMSGAQ KAA
Sbjct: 265 QMTAGSFCYIGPQGIVHGTTITVLNAGR-RMGL---DSLAGKVFVTAGLGGMSGAQPKAA 320

Query: 269 EIAKAVAIIAEVDQSRIKTRHSQGWISQIAESPEEALQLAQKAIDAKESTSIAYHGNIVD 328
           +IA  + +IAE+  + +  RH QGW+   ++  EE +   ++  + KE+ SI Y GN+VD
Sbjct: 321 KIAGCIGVIAEIISDTALLKRHQQGWLDVYSKDLEEIVNWIKEYREKKEAISIGYLGNVVD 380

Query: 329 LLE-YVNDKQIHVDLLSDQTSCHNVVDGGYCPVGISFDERTRLLARDKDTPHQMVDDTLA 387
           L E    + +  V+L SDQTS HN + GG+ P G++FD+  +++  D   P ++V ++L
Sbjct: 381 LWERLAEEPECLVELGSDQTSLHNDPLGGPVPAGLTFEQSNQMMTSDPVKPKKLVQNSLI 440

Query: 388 RHFEAIKTLTENGTYFFDYGNAFMKSVYDSGITEISKNGRNDKDGPIWPSYVEDIMGPML 447
           R   AI +   G YF+DYGNAF+     +G   + ++ ++DK  F +PSY++DIMG  +
Sbjct: 441 RQIAAIDKIAAKGMYFWDYGNAFLLECQRAGANLLKEDAQDDK-SFRYPSYMQDIMGD-I 498

Query: 448 FDYGYGPFRWVCLSGNHDDLVATDKAAMEAIDPDR--------RYQDRDNYNWIRDAEKN 499
           F  G+GPFRWVC SG  +DL  TD+ A + ID  +              + Q  DN  WI +AEKN
Sbjct: 499 FSMGFGPFRWVCTSGKPEDLRLTDQTACKIIDELKDIDVPEYVKQQYLDNKKWIEEAEKN 558

Query: 500 QLVVGTQARILYQDCIGRVTPIALKFNELVRKGKI-GFVMIGRDHHDVSGTDSPFRETSNI 558
           +LVVG+QARILY D  GRV +A  FNELV+ GK+   ++I RDHHDVSGTDSPFRETSN+
Sbjct: 559 KLVVGSQARILYSDRAGRVALASAFNELVKSGKVSAAIVISRDHHDVSGTDSPFRETSNV 618

Query: 559 KDGSNVTCDMAVQCTAGNAARGMSLVALHNGGGTGIGKAINGGFGLVILDGSRRIDEIIKS 618
            DGS  T DMAVQ   G++ RG + VALHNGGG G G  INGGFG+VLDGS       +
Sbjct: 619 YDGSAFTADMAVQCNCIGDSFRGATWVALHNGGGVGWGDVINGGFGIVLDGSSDAARRABG 678

Query: 619 AIAWDTMGGVARRNWARNEHAIETAIEYNRLHAGTDHITIPYLADDDLV 667
           + WD   GV RR+W+ N  A E AI+           +T+P  AD++L+
Sbjct: 679 MLNWDVPNGVTRRSWSGNAKQE-AIQRAEKQVDGLKVTLPVEADERLL 726
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2933

A DNA sequence (GASx2057) was identified in *S.pyogenes* <SEQ ID 8365> which encodes the amino acid sequence <SEQ ID 8366>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.1887(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAD35925 GB:AE001751
          formiminotransferase-
          cyclodeaminase/formiminotetrahydrofolate cyclodeaminase,
          putative [Thermotoga maritima]
 Identities = 160/296 (54%), Positives = 214/296 (72%), Gaps = 2/296 (0%)

Query: 3   KIVECIPNFSEGQNQAVIDGLVATAKSIPGVTLLDYSSDASHNRSVFTLVGDDQSIQEAA 62
           K++E +PNFSEG+ + V++ +VA AK    V +LD+S DA HNRSV TLVG+ +++  A
Sbjct: 2   KLIESVPNFSEGRRKEVVEKIVAEAKKYDRVHVLDMSMDADHNRSVITLVGEPENLINAL 61

Query: 63  FQLVKYASENIDMIKHHGEHPRMGATDVCPFVPIKDITQECVEISKQVAERINRELGIP 122
           F + K A+E ID+  H G+HPRMGA DV P VP+ + T +ECVE SK +  RI  ELGIP
Sbjct: 62  FDMTKKAASLIDLRNHTGQHPRMGAADVIPLVPLYNTTMEECVEYSKILGRRIGEELGIP 121

Query: 123 IPLYEDSATRPERQNLAKVRKGQPEGMPEKLLEEDWADYGDRKIHPTAGVTAVGARMPL 182
           ++LYE SATRPERQNLA +RKG+PEG  EK+ +  W PD+G  ++HPTAGVTAVGAR  L
Sbjct: 122 VYLYEKSATRPERQNLADIRKGEPEGPFRKIKDPLWKPDFGPDRVHPTAGVTAVGAREPL 181

Query: 183 VAFNVNLDTPNIDIAHKIAKIIRGSGGGYKYCKAIGVMLEDRHIAQVSMNMVNFEKCSLY 242
           +AFNVNL T ++ IA KIA+ IR S GG +Y KAIGV L+ R + QVS+N+ N +K  LY
Sbjct: 182 IAFNVNLGTRDVKIARKIARAIRFSSGGLRYVKAIGVDLKGRGVVQVSINITNHKKTPLY 241

Query: 243 RTFETIKFEARRYGVNVIGSEVIGLAPAKALIDVARYYLQVEDPDYHKQILEWHLL 298
           R FE IK EA RYGV V+GSE++GL P ++L+     YYL+ +      K++++E++LL
Sbjct: 242 RVFELIKMEAERYGVPVLGSEIVGLAPLKSLLKTVSYYLRTD--LNAKKVIESNLL 295
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2934

A DNA sequence (GASx2058) was identified in *S.pyogenes* <SEQ ID 8367> which encodes the amino acid sequence <SEQ ID 8368>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.2776(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA62653 GB:L33465 methenyl tetrahydrofolate cyclohydrolase
              [Methylobacterium extorquens]
 Identities = 79/198 (39%), Positives = 112/198 (55%)

Query: 7   SLTDFAKVLGSDAPAPGGGSAAALSGANGISLTKMVCELTLGKKKYADYQDITEIHAKS 66
            ++   F   L S AP PGGG AAA+SGA G +L  MVC LT+GKKKY + +   + ++   KS
Sbjct: 6   TIETFLDGLASSAPTPGGGGAAAISGAMGAALVSMVCNLTIGKKKYVEVEADLMQVLEKS 65

Query: 67  TALQASLLAAIDKDTEAFNLVSAVFDMPKNTDKDKAARRTAMQKALKTAAQSPFEMMTLM 126
              L+ +L   I  D EAF+ V   + +PK TDE+KAAR   +Q+ALKTA   P      +
Sbjct: 66  EGLRRTLTGMIADDVEAFDAVMGAYGLPKNTDEKKARAAKIQEALKTATDVPLACCRVC 125

Query: 127 VEALEITATAVGKSNTNAASDLGVAALNLKAGLQGAWLNVLINLSGIKDEDFVTDYRQKG 186
             E +++      K N N  SD GVA L+  AGL+ A LNV +N  G+ D  F  +  ++
Sbjct: 126 REVIDLAETVAEKGNLNVISDAGVAVLSAYAGLRSAALNVYVNAKGLDDRAFAEERLKEL 185

Query: 187 QALLDKGCHLADDIYTKI 204
            + LL +   L + IY  +
Sbjct: 186 EGLLAEAGALNERIYETV 203
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2935

A DNA sequence (GASx2061) was identified in *S.pyogenes* <SEQ ID 8369> which encodes the amino acid sequence <SEQ ID 8370>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3924(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2936

A DNA sequence (GASx2063) was identified in *S.pyogenes* <SEQ ID 8371> which encodes the amino acid sequence <SEQ ID 8372>. Analysis of this protein sequence reveals the following:

```
Possible site: 57

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -1.06   Transmembrane  231 - 247 ( 231 - 247)
   INTEGRAL    Likelihood = -0.53   Transmembrane    2 -  18 (   1 -  18)

----- Final Results -----
              bacterial membrane --- Certainty=0.1426(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB15971 GB:Z99124 histidase [Bacillus subtilis]
 Identities = 236/477 (49%), Positives = 321/477 (66%), Gaps = 2/477 (0%)

Query: 42  VINLDGESLTIEDVIAIARQGVACHIDDSAIRAVNASRKIVDDIVSERKVVVGVTTGFGS 101
           ++ LDG SLT  DV  +          + ++E V  SR  V+ IV +++ +YG+ TGFG
Sbjct: 1   MVTLDGSSLTTADVARVLFDFEEAAASEESMSRVVKSRAAVERIVRDEKTIYGINTGFGK 60

Query: 102 LCNVSISPEDTVQLQENLIRTHASGFGDPLPEDAVRAIMLIRINSLVKGYSGIRLSTIEK 161
             +V I  ED+  LQ NLI +HA G GDP PE   RA++L+R N+L+KG+SG+R   IE+
Sbjct: 61  FSDVLIQKEDSAALQLNLILSHACGVGDPFPECVSRAMLLLRANALLKGFSGVRAELIEQ 120

Query: 162 LLELNNKGVHPYIPEKGSLGASGDLAPLAHMVLPMLGLGKAYYKGELLSGQEALDKAGID 221
             LL  LNK VHP IP++GSLGASGDLAPL+H+ L ++G G+ +++GE +      L KAGI
Sbjct: 121 LLAFLNKRVHPVIPQQGSLGASGDLAPLSHLALALIGQGEVFFEGERMPAMTGLKKAGIQ 180

Query: 222 KISLAAKEGLALINGTTVLTAVGALATYDAIQLLKLSDLAGALSLEVHNGITSPFEENLH 281
             +L +KEGLALINGT  +TA+G +A  +A +L   +  +L++E  GI   F+E++H
Sbjct: 181 PVTLTSKEGLALINGTQAMTAMGVVAYIEAEKLAYQTERIASLTIEGLQGIIDAFDEDIH 240

Query: 282 TIRPQSGQLATARNIRNLLEGSQNTTVATQSRVQDPYTLRCMPQIHGASKDSIAYVKSKV 341
             R   Q+  A  IR  L  S  TT   + RVQD Y+LRC+PQ+HGA+  ++ YVK K+
Sbjct: 241 LARGYQEQIDVAERIRFYLSDSGLTTSQSELRVQDAYSLRCIPQVHGATWQTLGYVKEKL 300

Query: 342 DIEINSVTDNPIICKDG-EVISGGNFHGEPMAQPFDPLGIAISEIGNVSERRVERLVNSQ 400
           +IE+N+ TDNP+I  DG  VISGGNFHG+P+A   DPL IAISE+ N++ERR+ERLVN Q
Sbjct: 301 EIEMNAATDNPLIFNDGDKVISGGNFHGQPIAFAMDFLKIAISELANIAERRIERLVNPQ 360

Query: 401 LSKLPSFLVKYPGLNSGFMITQYACASLASENKVLAHPASVDSIPSCENQEDFVSMGTTA 460
           L+ LP FL  +PGL SG MI QYA ASL SENK LAHPASVDSIPS  NQED VSMGT A
Sbjct: 361 LNDLPPFLSPHPGLQSGAMIMQYAAASLVSENKTLAHPASVDSIPSSANQEDEVSMGTIA 420

Query: 461 ARKAFEILKNSRRIVATEIMAACQALDLKPENHELGKGTKVAYDLFREVNPIEHDK 517
           AR A++++ N+RR++A E + A QA++ +   H     TK  +  RK V  I+ D+
Sbjct: 421 ARHAYQVIANTRRVIAIEAICALQAVEYRGIEH-AASYTKQLFQEMRKVVPSIQQDR 476
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2937

A DNA sequence (GASx2064) was identified in *S.pyogenes* <SEQ ID 8373> which encodes the amino acid sequence <SEQ ID 8374>. Analysis of this protein sequence reveals the following:

```
Possible site: 44

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4483(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAG06563 GB:AE004741 probable arginase family protein
          [Pseudomonas aeruginosa]
 Identities = 99/275 (36%), Positives = 147/275 (53%), Gaps = 9/275 (3%)
```

```
Query: 53   LIGFKSDKGVYINNGRVGAVESPAAIRTQLAKFPWHLSNQVMVYDVGNIDGNRSLEQLQ 112
            L+GF SD+GV  N GR GA   P A+R  LA   WH G Q  +YD G+I  +  LE  Q
Sbjct: 42   LLGPASDEGVRRNQGRQGARHGPPALRRALANLAWH-GEQA-IYDAGDIVAGD-DLEAAQ 98

Query: 113  NSLSKAIKRMCDLNLKPIVLGGGHETAYGHYILGLRQSHSPSDDL---AVINMDAHFDLRP 169
             ++ +  +     + + LGGGHE AY  + GL + LS  + L    ++N DAHFDLR
Sbjct: 99   ECYAQRVADLLACGHRVVGLGGGHEIAYASFAGLARHLRRIGIINFDAHFDLRH 158

Query: 170  YDQIGPNSGTGPRQMFDDAVADKRLFKYFVLGIQEHNNNLFLPDFVAKSKGIQFLTGQDI 229
             ++   +SGT PRQ+ +  A    F Y  LGI   +N  LFD  A+  G+++L  + +
Sbjct: 159  AERA--SSGTPPRQIAKLCQASWPPFAYCCLGISRLSNTAALFD-QAQRLGVRYLLDRQL 215

Query: 230  YQMGHQKVCRAIDRFLEGQERVVLFTDMDCFSVGAAPGVSAIQSLGVDPNLAVLVLQHIA 289
             ++   +D FL+  + +YLT+ +D      APGVSA  + GV+  +   +++
Sbjct: 216  QPWNLERSRAPLDGFLQSVDHLYLTVCLDVLPAAQAPGVSAPSAHGVEMPVVEHLVRRAK 275

Query: 290  ASGKLVGFDVVSVSPPHDIDNHTANLAATFIFYLW 324
             ASGKL   D+ E++P  D D  TA +AA  +  LV
Sbjct: 276  ASGKLRLADIAELNPQLDSDQRTARIAARLVDSLV 310
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2938

A DNA sequence (GASx2065R) was identified in *S.pyogenes* <SEQ ID 8375> which encodes the amino acid sequence <SEQ ID 8376>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.37   Transmembrane  375 - 391 ( 375 - 392)

----- Final Results -----
               bacterial membrane --- Certainty=0.1150(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB37582 GB:AL035569 putative regulatory protein [Streptomyces
      coelicolor A3(2)]
 Identities = 95/437 (21%), Positives = 177/437 (39%), Gaps = 28/437 (6%)

Query: 271  EVGALLLIGDTGIGKRTLARQVLANQTQTFQIVTAKCFREEAMDSL--LPWRNILDGLGD 328
             E  ALLL G+ G+GK  L  +  A   +V      E   D L   P+   L  L
Sbjct: 95   EPQALLLGGEAGVGKTRLVEEPAAAADRRGAVVALGGCVRIGADGLPFAPFSTALRALRR 154

Query: 329  LVIQNRLLLTTKAWKAALKRCFP-VATIFQEDNNQPFIKDHTSLLVSPIVDILQHLAEIKA 387
              + +          +  L R  P +A            ++  + L      +L+ +A
Sbjct: 155  HLPEELAAAAAQQEEELARLLPKLABGTPVTGGGRHDRESMARLFELTARLLERVAARHT 214

Query: 388  LVILIEDCHWMDEDSLTLLQRVMNQLVHYPIAFVLT-------KHLGTTPELGLCLNALM 440
             +V++ED HW D  +  L+  ++  L    +  +T      +   P L  L+ L
Sbjct: 215  VVLVLEDLHWADASTRHLIAYLARTLRTGRLVVLATYRSDDIHRRHPLRPLIAE-LDRLR 273

Query: 441  SQGRLESICLEPFNRQESLVYINSQLGSQPVTAEEMEHLYQASQGNDPFLSEYTQALLKH 500
             + RLE   L  F R E    I   L  +P    +++ +++ S GN FF+ E   A  R
Sbjct: 274  TVRRLE---LGRPTRDEWGRQIAGILAHEP-DQLVDEIFERSDGNAFFVRELAVA-ARV 328

Query: 501  EKFVPLTPAIKAKLGLKLANLSSRDDALLNYLSCCRRPIPLNTLAQLMLLPLEEVIEMVD 560
              LT +++  L +++  L      +  ++      +   LA +  L  +++IE +
Sbjct: 329  GSCTGLTDSLRDLLLVRVEALPESAQRVARIVAVGGSTVEYRLLAAVARLAEDDLIEALR 388
```

```
Query: 561 NLGHYYILVKESVGKEVLISFRQRIIQLYSYDRLSLSKRRLLHQQIAKRLGDLLPILTPS 620
           +  +  IL+    G+      FR  +++     D L    +R  L+ + A+ L D   P L P+
Sbjct: 389 SAVNANILLPAPDGDG--YRFRHSLVREAVGDDLLPGERSRLNRRYAEAL-DADPTLVPA 445

Query: 621 PHLLDDIAYHYQSSRQVIKALKYNLNYLDATLPFQHKLFTYSKGIGSLEKGDRDHQRLM 680
             +A ++  +    KAL     LDA++  +    YS+ +  LE++          L
Sbjct: 446 AERVMRLASIYWHAHAPAKALP---AVLDASVEARRR--HAYSEQLRLERA----MELW 496

Query: 681 EEQFDKIRQSIADLRLT 697
           +   D +R ++  ++ T
Sbjct: 497 DSAPDDVRATLRPVDCT 513
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2939

A DNA sequence (GASx2072) was identified in *S.pyogenes* <SEQ ID 8377> which encodes the amino acid sequence <SEQ ID 8378>. Analysis of this protein sequence reveals the following:

```
Possible site: 14

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3702(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2940

A DNA sequence (GASx2074R) was identified in *S.pyogenes* <SEQ ID 8379> which encodes the amino acid sequence <SEQ ID 8380>. Analysis of this protein sequence reveals the following:

```
Possible site: 37

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.90   Transmembrane    21 -  37 (  21 -  38)

----- Final Results -----
               bacterial membrane --- Certainty=0.1362(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2941

A DNA sequence (GASx2075R) was identified in *S.pyogenes* <SEQ ID 8381> which encodes the amino acid sequence <SEQ ID 8382>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3545(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2942

A DNA sequence (GASx2076R) was identified in *S.pyogenes* <SEQ ID 8383> which encodes the amino acid sequence <SEQ ID 8384>. Analysis of this protein sequence reveals the following:

```
Possible site: 34

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2340(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAC44494 GB:U44893 orf108; unknown function [Butyrivibrio
          fibrisolvens]
 Identities = 42/75 (56%), Positives = 55/75 (73%)

Query: 1   LLKGTLRFGQLKSSIGSVSQKVLTAQLRAMEADGLVHREVVAEVPPRVEYSLTETGLSLA 60
           LL    RF +LK++  +SQKVLT  LR+ME DG++ R VV EVPPRVEYSL+E G S+
Sbjct: 31  LLVRPWRFNELKNNLEGISQKVLTDSLRSMERDGIITRTVYPEVPPRVEYSLSELGESMR 90

Query: 61  PVIEAMSDWGQTYQR 75
           P+I+AM  WG  Y+E
Sbjct: 91  PIIKAMEQWGTEYKE 105
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2943

A DNA sequence (GASx2097) was identified in *S.pyogenes* <SEQ ID 8385> which encodes the amino acid sequence <SEQ ID 8386>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -3.40   Transmembrane    26 -  42 (  23 -  44)
```

```
----- Final Results -----
                bacterial membrane --- Certainty=0.2359(Affirmative) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2944

A DNA sequence (GASx2098) was identified in *S.pyogenes* <SEQ ID 8387> which encodes the amino acid sequence <SEQ ID 8388>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.1385(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2945

A DNA sequence (GASx2100) was identified in *S.pyogenes* <SEQ ID 8389> which encodes the amino acid sequence <SEQ ID 8390>. Analysis of this protein sequence reveals the following:

```
Possible site: 23

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2138(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAA98589 GB:L44593 ORF79; putative [Lactococcus lactis phage
             BK5-T]
 Identities = 34/62 (54%), Positives = 44/62 (70%)

Query: 3   QITLKAARINAGYTLKQVAGAVGKNPQTISKYEKDGSDISLGLLQKLSSLYGVTIDNLFL 62
           +I LKAAR NA ++ K+VA  VGKN QTI  YEKDGS++I + L  KL+ +Y   ID +FL
Sbjct: 8   KILKLKAARTNADPSAKEVAEIVGKNYQTILSYEKDGSTEIPMSLAIKLAEIYDYPIDFIFL 67

Query: 63 GK 64
             GK
Sbjct: 68 GK 69
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2946

A DNA sequence (GASx2103) was identified in *S.pyogenes* <SEQ ID 8391> which encodes the amino acid sequence <SEQ ID 8392>. Analysis of this protein sequence reveals the following:

```
Possible site: 39

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3316(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2947

A DNA sequence (GASx2104) was identified in *S.pyogenes* <SEQ ID 8393> which encodes the amino acid sequence <SEQ ID 8394>. Analysis of this protein sequence reveals the following:

```
Possible site: 55

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4371(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2948

A DNA sequence (GASx2105) was identified in *S.pyogenes* <SEQ ID 8395> which encodes the amino acid sequence <SEQ ID 8396>. Analysis of this protein sequence reveals the following:

```
Possible site: 40

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2263(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2949

A DNA sequence (GASx2106) was identified in *S.pyogenes* <SEQ ID 8397> which encodes the amino acid sequence <SEQ ID 8398>. Analysis of this protein sequence reveals the following:

```
Possible site: 32

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -6.42   Transmembrane    9 -  25 (   6 -  29)

----- Final Results -----
               bacterial membrane --- Certainty=0.3569(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2950

A DNA sequence (GASx2107) was identified in *S.pyogenes* <SEQ ID 8399> which encodes the amino acid sequence <SEQ ID 8400>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1355(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2951

A DNA sequence (GASx2108) was identified in *S.pyogenes* <SEQ ID 8401> which encodes the amino acid sequence <SEQ ID 8402>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3050(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2952

A DNA sequence (GASx2109) was identified in *S.pyogenes* <SEQ ID 8403> which encodes the amino acid sequence <SEQ ID 8404>. Analysis of this protein sequence reveals the following:

```
Possible site: 13

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3628(Affirmative) < succ>
            bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB46557 GB:AJ242479 putative replication protein [Streptococcus thermophilus]
 Identities = 143/242 (59%), Positives = 180/242 (74%), Gaps = 2/242 (0%)

Query: 1   MAIYEARGFSSYLY--PYKGPLEPPDYIAQFRPLKPPEDIDIEEYKRTQAPYCLSGKVTA 58
           MAIYE+RGF + L+         +PF ++A FRP+K P+  DI ++KR  APYC+SG+V
Sbjct: 1   MAIYESRGFGNILHLNNSNASKDPFKFVATFRPMKVPQGEDIADFKRYHAPYCISGEVKQ 60

Query: 59  EKNGSYKRNNASLVYRDLIFLDYDEIETGVNLPKIVSQTLWEYSYIIYPTIKHTPEKPRY 118
           +++G+YKRNNASL+YRDLIFLDYD++E   + P+ VS  L  YSY+IYPTIKHT EKPRY
Sbjct: 61  DEDGNYKRNNASLLYRDLIFLDYDKLEASTDFPRAVSNALNGYSYVIYPTIKHTAEKPRY 120

Query: 119 RLVMKPSDVWTEATYKQVVKEIADKIGLPFDLASLTWSQLQGLPVTTGDPEDYQRYVNHG 178
           RLV+KP+D M E TYK    +EIADKIGLPFD +SLTWSQLQGLPVTTGDPE Y+R VN G
Sbjct: 121 RLVVKPTDKMDEQTYKATAQEIADKIGLPFDDSSLTWSQLQGLPVTTGDPEKYERIVNRG 180

Query: 179 LDYPVPKNGSTPNRQVVTTYTPRPRSQRSITMRVIDTLFNGPGNNGGRNVALTKFVGLLP 238
           YPV    +        +TPR    +S+TMRV+DTL NGPG+EGGRN+ +T+FVGLL
Sbjct: 181 RCYPVANPNTVKANHSPNYHTPRQSGDKSLTMRVVDTLLNGPGDEGGRNIEVTRFVGLLL 240

Query: 239 NK 240
           +K
Sbjct: 241 SK 242
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2953

A DNA sequence (GASx2110) was identified in *S.pyogenes* <SEQ ID 8405> which encodes the amino acid sequence <SEQ ID 8406>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.5215(Affirmative) < succ>
```

```
bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAB46558 GB:AJ242479 putative DNA primase [Streptococcus thermophilus]
 Identities = 274/548 (50%), Positives = 363/548 (66%), Gaps = 17/548 (3%)

Query: 17  DLKNLENRITEARE------NRDKYFSTFKGVRQQLIKRCQRMKDEAFKIAYDGVMADSK 70
           DL  LE E  R++++       +ED Y  TFK +R Q I   ++ K+ A++  YD  M + K
Sbjct: 8   DLTKLEEEYNESKKEASTLFDEDGYLKTFKDIRKQFINILEQKKEIAYQKGYDLYMNNPK 67

Query: 71  HLENVKAGKLTEVQHE-------ELAKEKGQERASEKALPKTPLGVAIMLKHYLRFIRVKP 123
            L  +      E   E       E AK++G++A + A PKTPL  A   LK Y+RFIR++P
Sbjct: 68  VLLKLAKAEKDEENGELIRKTVIEDAKKEGEKAKKNATPKTPLECAEFLKKYIRFIRIRP 127

Query: 124 EAQGQKAPLYFFHPDHGVWLEDNEFLQDLISVIFPNATEKQAFDTLYKIARQSQLKEIQR 183
            + +G++     F     G++LED+EFL DL+  I PN TE+   D LYKIA    LK+ Q
Sbjct: 128 KGKGRERLYTFTRQILGIYLEDDEFLHDLMVTIHPNNTERLGNDALYKIAHSVPLKDKQE 187

Query: 184 EYTVIGNQLYNYKTGQFEELTPDITVTRKIKTGYNKKAKEPTIKGWKFTAWLLELPDGDA 243
            Y V+G +LYN +TG+F +   P I VTRK++ GYN  A EP I GWKFT WL  LF+GD
Sbjct: 188 NYVVVGGELYNNETGEFTQPDPRIIVTRKVRWGYNEDATEPIIDGWKFTWLKGLFNGDR 247

Query: 244 ELYNLAIQIIKASTGQSLQKIFWLPGEGGTGKGTFQQLLINLVGMDNVASLKITELAKS 303
            + Y+LAIQII+A+ITG++L+ IFWL+GEGGTGKGTFQ LL NLVG +NVAS KI + A
Sbjct: 248 DSYDLAIQIIRATITGKTLENIFWLYGEGGTGKGTFQTLLENLVGSENVASFKI-DGASG 306

Query: 304 RFTTSILLGKSIVIGDDIQKDAVIKDTSDIFSLATGDIMTIEDKGKRPYSIRLNMTVVQS 363
            +F TSIL+GK++VIGDDIQKD VIKDTS +FSLATGD + IEDKGKRPY+ R  MTVVQS
Sbjct: 307 KFDTSILIGKTVVIGDDIQKDVVIKDTSVVFSLATGDPIRIEDKGKRPYTTRKRMTVVQS 366

Query: 364 SNGLPRMNGDKSAIDRRFRILPFTKVFKGKPNKAIRNDYINRKEVLEYLLKLAIETPITD 423
            SNG PRMN D+ AI+RRFR+L F+++ KGK +K I+NDY+ RKEVLEY +KLAIETP  D
Sbjct: 367 SNGPFRMNADQKAINRRFRVLYFSEL-KGKADKRIKNDYVGRKEVLEYFVKLAIETFFRD 425

Query: 424 INPKASIEILEEHHKEMNPVIDFVSKPFTDE-LTSEFIPNSFVYHVWKGPLEYYDIKQ-I 481
            +NP+ SIE L+E +KEMNPV DFV +FF DE +   ++PN +V+  +K + E    +
Sbjct: 426 VNPQKSIEFLDEAYKEMNPVADFVDRFFNDEVIKCNYVPNGYVPECFKAYCEKNQNRNYF 485

Query: 482 KSERGLHKSIKSNLPEGFEAGQKVIPVGRQLHTGFYPKEDLPLFASASYANGRASPEKRK 541
             + R LHK+IK  LP+ F   +  I  G++ +  F P       +   +Y NGR   E ++
Sbjct: 486 LNSRTLHKQIKKILPKTFRPKEVTIKKGQKFYREFNPHLVSNPWHFDAYDNGRNKKEDQQ 545

Query: 542 KPKNERGY 549
             K ERGY
Sbjct: 546 DAKKERGY 553
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2954

A DNA sequence (GASx2111) was identified in *S.pyogenes* <SEQ ID 8407> which encodes the amino acid sequence <SEQ ID 8408>. Analysis of this protein sequence reveals the following:

```
Possible site: 41

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.0994(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2955

A DNA sequence (GASx2112) was identified in *S.pyogenes* <SEQ ID 8409> which encodes the amino acid sequence <SEQ ID 8410>. Analysis of this protein sequence reveals the following:

```
Possible site: 54

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3058(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2956

A DNA sequence (GASx2114) was identified in *S.pyogenes* <SEQ ID 8411> which encodes the amino acid sequence <SEQ ID 8412>. Analysis of this protein sequence reveals the following:

```
Possible site: 37

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.2815(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2957

A DNA sequence (GASx2115R) was identified in *S.pyogenes* <SEQ ID 8413> which encodes the amino acid sequence <SEQ ID 8414>. Analysis of this protein sequence reveals the following:

```
Possible site: 27

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
              bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial outside --- Certainty=0.0000(Not Clear) < succ>
bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2958

A DNA sequence (GASx2116) was identified in *S.pyogenes* <SEQ ID 8415> which encodes the amino acid sequence <SEQ ID 8416>. Analysis of this protein sequence reveals the following:

```
Possible site: 56

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.4213(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2959

A DNA sequence (GASx2117) was identified in *S.pyogenes* <SEQ ID 8417> which encodes the amino acid sequence <SEQ ID 8418>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3091(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2960

A DNA sequence (GASx2118) was identified in *S.pyogenes* <SEQ ID 8419> which encodes the amino acid sequence <SEQ ID 8420>. Analysis of this protein sequence reveals the following:

```
Possible site: 41

>>> Seems to have an uncleavable N-term signal seq
```

```
----- Final Results -----
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
               bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2961

A DNA sequence (GASx2119) was identified in *S.pyogenes* <SEQ ID 8421> which encodes the amino acid sequence <SEQ ID 8422>. Analysis of this protein sequence reveals the following:

```
Possible site: 22

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2531(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF63071 GB:AF158600 gp137 [Streptococcus thermophilus
          bacteriophage Sfi11]
 Identities = 41/121 (33%), Positives = 65/121 (52%), Gaps = 3/121 (2%)

Query: 4   KNAIRKLKEFHRWQRIAN-SLDLTYTELYQFDIEYHPTRR--KHLEISRECALEELDAIR 60
           K   RKL+E+ RW+ IA+ S +    T+ + F        +  +++ + R  AL EL+AI
Sbjct: 13  KRCKRKLREYPRWRETAHDSAEQKITQEPTFMPRGGGVNKPVENIAVRRVDALNELEAIE 72

Query: 61  YAINQLSKVEYRQILIECYLISEEKTQQDIMELNGSQSWTYESKKRALLEFVEFYRDGAL 121
            A+N L + +YR+ILIE YL     K    I + +   ++ + E    ++L F E YRDG L
Sbjct: 73  QAVNGLYRPDYRRILIEKYLAYPPKPNWQIAQSIGFERTAFQELLNNSILAFAELYRDGRL 133
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2962

A DNA sequence (GASx2120) was identified in *S.pyogenes* <SEQ ID 8423> which encodes the amino acid sequence <SEQ ID 8424>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
               bacterial cytoplasm --- Certainty=0.2666(Affirmative) < succ>
                bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                 bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2963

A DNA sequence (GASx2121) was identified in *S.pyogenes* <SEQ ID 8425> which encodes the amino acid sequence <SEQ ID 8426>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have a cleavable N-term signal seq.

----- Final Results -----
            bacterial outside --- Certainty=0.3000(Affirmative) < succ>
           bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
          bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2964

A DNA sequence (GASx2123R) was identified in *S.pyogenes* <SEQ ID 8427> which encodes the amino acid sequence <SEQ ID 8428>. Analysis of this protein sequence reveals the following:

```
Possible site: 21

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
          bacterial cytoplasm --- Certainty=0.3441(Affirmative) < succ>
           bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2965

A DNA sequence (GASx2132) was identified in *S.pyogenes* <SEQ ID 8429> which encodes the amino acid sequence <SEQ ID 8430>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
           bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
            bacterial outside --- Certainty=0.0000(Not Clear) < succ>
          bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2966

A DNA sequence (GASx2136) was identified in *S.pyogenes* <SEQ ID 8431> which encodes the amino acid sequence <SEQ ID 8432>. Analysis of this protein sequence reveals the following:

```
Possible site: 30

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -3.19   Transmembrane   57 -  73 (  54 -  78)

----- Final Results -----
               bacterial membrane --- Certainty=0.2275(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAB18271 GB:U74623 CadX [Staphylococcus lugdunensis]
 Identities = 50/110 (45%), Positives = 76/110 (68%)

Query: 11  MKKDSICQVGVINQQNVTTATNYLSKEKVQKSLRILSKFTDNKQINIIFYLLAVEELCVC 70
           M  ++ C V  +++  V  A ++LE +K +K L IL K  D K++ II  L+  +ELCVC
Sbjct: 1   MSYENACDVICVHEDKVNNALSFLEDDKSKKLLNILEKICDEKKLKIILSLIKEDELCVC 60

Query: 71  DIACLLNLSMASASHHLRKLANQNILDTRREGKIIYYFIKDEEIRDFFNQ 120
           DI+ +L +S+AS SHHLR L   ++LD  ++GK+ YYFIKD+EIR+FF++
Sbjct: 61  DISLILKMSVASTSHHLRLLYKNDVLDFYKKGKMAYYFIKDDEIREFFSK 110
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2967

A DNA sequence (GASx2137) was identified in *S.pyogenes* <SEQ ID 8433> which encodes the amino acid sequence <SEQ ID 8434>. Analysis of this protein sequence reveals the following:

```
Possible site: 49

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4582(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2968

A DNA sequence (GASx2139) was identified in *S.pyogenes* <SEQ ID 8435> which encodes the amino acid sequence <SEQ ID 8436>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood = -5.89   Transmembrane   63 -  79 (  54 -  80)

----- Final Results -----
               bacterial membrane --- Certainty=0.3357(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2969

A DNA sequence (GASx2141R) was identified in *S.pyogenes* <SEQ ID 8437> which encodes the amino acid sequence <SEQ ID 8438>. Analysis of this protein sequence reveals the following:

```
Possible site: 19

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4663(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2970

A DNA sequence (GASx2142) was identified in *S.pyogenes* <SEQ ID 8439> which encodes the amino acid sequence <SEQ ID 8440>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have a cleavable N-term signal seq.
   INTEGRAL    Likelihood =-10.08   Transmembrane  143 - 159 ( 135 - 165)
   INTEGRAL    Likelihood = -7.64   Transmembrane   53 -  69 (  49 -  79)
   INTEGRAL    Likelihood = -7.17   Transmembrane  252 - 268 ( 248 - 275)
   INTEGRAL    Likelihood = -6.74   Transmembrane  186 - 202 ( 183 - 208)
   INTEGRAL    Likelihood = -5.63   Transmembrane  220 - 236 ( 218 - 240)
   INTEGRAL    Likelihood = -5.26   Transmembrane  116 - 132 ( 115 - 136)
   INTEGRAL    Likelihood = -2.02   Transmembrane   85 - 101 (  85 - 101)
   INTEGRAL    Likelihood = -0.64   Transmembrane  165 - 181 ( 165 - 181)

----- Final Results -----
               bacterial membrane --- Certainty=0.5034(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

```
bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAD35257 GB:AE001701 conserved hypothetical protein [Thermotoga maritima]
 Identities = 81/275 (29%), Positives = 137/275 (49%), Gaps = 29/275 (10%)

Query: 9   FKGMIIALGFILPGVSGSVLAAILGIYERKMISFLAHMRDNFIRNVLFFLPVGIG---GIL 65
           F G++ +  ++PGVSGG +A ++G+YR++I  +      ++  +PVG G   G+
Sbjct: 7   FSGVLMGIANVVPGVSGGTIAVLMGVYEKLIRSVNSFPHGNSRSLKVLIPVGAGVLVGVF 66

Query: 66  GIALFSPFVRFLLKHYQVSVLWGFAGAIVGTIPSLIKESTKQSQRDKADWLWLVLTFVIS 125
           GIA F   +E  L  Y V   + F G I   I S +K  TK+    K     + + FV+
Sbjct: 67  GIARF---LEIFLSKYPVPTHPFFLGLI---IVSFVK--TKEYFSIKP----VNIFFVLL 114

Query: 126 GLGLYFLNDLIG--TLPANFLTFILAGALIALGVLVPGLSPSNLLLILGLYGPMLIGPFS 183
           G+ L F+    G  T   +    +L G + A  ++VPG+S S +LLI G+Y  +L
Sbjct: 115 GMFLIFMLHFSGETTAKESMFLLVLGGFVAATAMVVPGISGSLLLLIPGVYDHVLYLVSH 174

Query: 184 LDLLGTFLPIAIGGVLAILAFSKSMDYALQHHHSKVYHFIGIVLSSTLLILIPNSSSPE 243
           L ++G  L  +IG V  IL   K M++ L+    + Y FI G++L+S L  ++P     +
Sbjct: 175 L-IIGELLIFSIGVVAGILVSVKIMNFLLKRFRETSYSFIGGMILAS-LVEVLPKGMYTN 232

Query: 244 SISYSHAGILTWLMAPVLFALGIWLGLWMSQLEEK 278
            +          L + +   L + LG ++  +E+K
Sbjct: 233 VV----------LPSVLSLVLSLTLGFFLLYIEKK 257
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2971

A DNA sequence (GASx2143R) was identified in *S.pyogenes* <SEQ ID 8441> which encodes the amino acid sequence <SEQ ID 8442>. Analysis of this protein sequence reveals the following:

```
Possible site: 20

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
            bacterial cytoplasm --- Certainty=0.3964(Affirmative) < succ>
             bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
              bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae*.

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:HAB05000 GB:AP001511 unknown conserved protein in others
           [Bacillus halodurans]
 Identities = 28/78 (35%), Positives = 37/78 (46%)

Query: 44  EVDKVFIVPLRQTLFTDFVYYRLEVTPIETTDFFFDRIRNGKYYQPSQEYRSIPFYENLE 103
           EVD VF VP+   +   P  YR+ V       FF +RI N   YQ S    +  FY
Sbjct: 127 EVDHVFTVPIDHFISHPPRQYRINVHFEPGAGFPIERIANQSAYQKSTRQITESFYYQS 186

Query: 104 ETIWGMTAQFTKCLITDIL 121
             IWG+TA+  +  IL
Sbjct: 187 YVIWGLTAKILRHVITIL 204
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2972

A DNA sequence (GASx2144R) was identified in *S.pyogenes* <SEQ ID 8443> which encodes the amino acid sequence <SEQ ID 8444>. Analysis of this protein sequence reveals the following:

```
Possible site: 17

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4761(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2973

A DNA sequence (GASx2145) was identified in *S.pyogenes* <SEQ ID 8445> which encodes the amino acid sequence <SEQ ID 8446>. Analysis of this protein sequence reveals the following:

```
Possible site: 25

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -4.09   Transmembrane    2 -  18 (   1 -  19)

----- Final Results -----
              bacterial membrane --- Certainty=0.2635(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA49519 GB:X69895 X [Bacillus sphaericus]
 Identities = 40/97 (41%), Positives = 57/97 (58%), Gaps = 5/97 (5%)

Query: 10  IEFLILAIVEKNDSYGYDISQTIKLVAN----IKESTLYPILKKLEKAGFLTTYSQE-HQ 64
           ++ +IL ++ + D YGY+ISQ I    N    IKE+TLY + ++LEK   +  Y  +
Sbjct: 11  LDSIILRLLESKDRYGYEISQEISNRTNNSPQIKEATLYAVFQRLSKKEVIEAYYGDVSD 70

Query: 65  GKKRKYYAVTSSGRAQLIFLKKEWQSYKFALDGIIEG 101
           G KRKYY +TS G+A L  L KEW   K  +D  +EG
Sbjct: 71  GGKRKYYRITSLGKAYLSELVKEWAEVKEIDLFMEG 107
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2974

A DNA sequence (GASx2146) was identified in *S.pyogenes* <SEQ ID 8447> which encodes the amino acid sequence <SEQ ID 8448>. Analysis of this protein sequence reveals the following:

```
Possible site: 56

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-14.75   Transmembrane   97 - 113 (  77 - 143)
```

```
INTEGRAL    Likelihood = -6.85   Transmembrane  116 - 132 ( 114 - 143)
INTEGRAL    Likelihood = -5.68   Transmembrane  156 - 172 ( 149 - 175)
INTEGRAL    Likelihood = -5.47   Transmembrane   79 -  95 (  77 -  96)

----- Final Results -----
              bacterial membrane --- Certainty=0.6901(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2975

A DNA sequence (GASx2147) was identified in *S.pyogenes* <SEQ ID 8449> which encodes the amino acid sequence <SEQ ID 8450>. Analysis of this protein sequence reveals the following:

```
Possible site: 31

>>> Seems to have an uncleavable N-term signal seq
INTEGRAL    Likelihood = -7.11   Transmembrane    8 -  24 (   6 -  30)

----- Final Results -----
              bacterial membrane --- Certainty=0.3845(Affirmative) < succ>
               bacterial outside --- Certainty=0.0000(Not Clear) < succ>
             bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:AAF04457 GB:AF078161 lacunin [Manduca sexta]
 Identities = 68/310 (21%), Positives = 117/310 (36%), Gaps = 12/310 (3%)

Query: 55   DIDSSASTITVETGPVQRPTVTYYTHPKLIDPIVTTVTGKTLSLSQTPKDVVITGGIEIL 114
             DI+ + ++ + E+    T++  T     +    TT T  T  +S T +    I      +
Sbjct: 1004 DIEGTTASGSTESTFTDETTMSKVTESSSVAESETTKTITTEEVSGTSESASINSDKTTM 1063

Query: 115  GFTLNNSRQEKNYRSIT--ITVPEKTSLNEVKASNVPHTTLSNLT--VQDMQFDGNLTLL 170
                 ++ +      IT  +TV E+TS           TT+S ++   +        T
Sbjct: 1064 TTLSEDTGKTSVSERITTEMTVTEETSETSPTEGTSDKTTMSTVSETESSSSVTEKTTTE 1123

Query: 171  HTKVKKATITGMLEATKSQLTNLELKADYSPSNLTDSSVE-NGTISLGNSQLTTKDTTLK 229
              T V+ AT     E T S  T +   ++ S      +++ E   T  +      T   T+ K
Sbjct: 1124 TTVVENATDISSTEVTASDKTTMTTMSEESEKTTEEATTEITTVTKEVTESSSSTETATSDK 1183

Query: 230  AVNIQSLHPGGIE-AERTTLENVTFTVSKSKKEEENDYYDNDAIPTAHAILTLKGTNTTTG 288
             ++  S   G    AE +T E VT T   + EE                T+  +T+K T T
Sbjct: 1184 TISTLSEETGKTSVAEESTTEKVTETVTTMPEETGK------TITSEEITIKTTVTEEP 1237

Query: 289  GDIDVDITLTKAKAIAYRARTENGKVSLGSQLTPAKICKESTSDVISVVAENKAATGNLT 348
              D+     +T  K     A  E GK S+  + T       E++++  S  A     T   T
Sbjct: 1238 TDVGSSEAITSDKTTVSTASEETGKYSVSEEETVKTTVAEASTPSSTEAITSDKTKMST 1297

Query: 349  VNLNKGDITI 358
             ++   G  ++
Sbjct: 1298 ISEETGKTSV 1307
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2976

A DNA sequence (GASx2148R) was identified in *S.pyogenes* <SEQ ID 8451> which encodes the amino acid sequence <SEQ ID 8452>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have an uncleavable N-term signal seq

----- Final Results -----
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2977

A DNA sequence (GASx2160) was identified in *S.pyogenes* <SEQ ID 8453> which encodes the amino acid sequence <SEQ ID 8454>. Analysis of this protein sequence reveals the following:

```
Possible site: 29

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.1630(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2978

A DNA sequence (GASx2170R) was identified in *S.pyogenes* <SEQ ID 8455> which encodes the amino acid sequence <SEQ ID 8456>. Analysis of this protein sequence reveals the following:

```
Possible site: 37

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood =-13.32   Transmembrane  181 - 197 ( 175 - 203)

----- Final Results -----
               bacterial membrane --- Certainty=0.6328(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2979

A DNA sequence (GASx2174) was identified in *S.pyogenes* <SEQ ID 8457> which encodes the amino acid sequence <SEQ ID 8458>. Analysis of this protein sequence reveals the following:

```
Possible site: 28

>>> Seems to have an uncleavable N-term signal seq
   INTEGRAL    Likelihood = -2.39   Transmembrane    3 -  19 (   3 -  19)

----- Final Results -----
               bacterial membrane --- Certainty=0.1956(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2980

A DNA sequence (GASx2181R) was identified in *S.pyogenes* <SEQ ID 8459> which encodes the amino acid sequence <SEQ ID 8460>. Analysis of this protein sequence reveals the following:

```
Possible site: 24

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3751(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2981

A DNA sequence (GASx2185R) was identified in *S.pyogenes* <SEQ ID 8461> which encodes the amino acid sequence <SEQ ID 8462>. Analysis of this protein sequence reveals the following:

```
Possible site: 26

>>> Seems to have no N-terminal signal sequence
   INTEGRAL    Likelihood = -0.90   Transmembrane   18 -  34 (  18 -  34)

----- Final Results -----
               bacterial membrane --- Certainty=0.1362(Affirmative) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
              bacterial cytoplasm --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has no significant homology with any sequences in the GENPEPT database.

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2982

A DNA sequence (GASx2186R) was identified in *S.pyogenes* <SEQ ID 8463> which encodes the amino acid sequence <SEQ ID 8464>. Analysis of this protein sequence reveals the following:

```
Possible site: 61

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.4803(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA78948 GB:Z17279 transposase [Streptococcus salivarius]
 Identities = 48/77 (62%), Positives = 55/77 (71%), Gaps = 1/77 (1%)

Query: 1   VSMKPIDLSKMVSIRKRSKKVMKTNKKTLGKSIRERPEYINDRSEPGHWEIDLALGKKTK 60
           + +K IDL + V IRK+  K    T KK LGKSIRERPE IN+RS PG WEID  LG KT
Sbjct: 150 LEIKVIDLPRAVRIRKKFTKRPST-KKHLGKSIRERPEINNRSRPGDWEIDSVLGKKTI 208

Query: 61  SEAVMLTLVERQTRYAL 77
            E  +LTLVERQTRYA+
Sbjct: 209 GEPSILTLVERQTRYAV 225
```

Based on this analysis, it was predicted that this GAS-specific protein and its epitopes, could be useful antigens for vaccines or diagnostics.

Example 2983

A DNA sequence (GASx2187R) was identified in *S.pyogenes* <SEQ ID 8465> which encodes the amino acid sequence <SEQ ID 8466>. Analysis of this protein sequence reveals the following:

```
Possible site: 50

>>> Seems to have no N-terminal signal sequence

----- Final Results -----
              bacterial cytoplasm --- Certainty=0.3287(Affirmative) < succ>
               bacterial membrane --- Certainty=0.0000(Not Clear) < succ>
                bacterial outside --- Certainty=0.0000(Not Clear) < succ>
```

No corresponding DNA sequence was identified in *S.agalactiae.*

The protein has homology with the following sequences in the GENPEPT database:

```
>GP:CAA78948 GB:Z17279 transposase [Streptococcus salivarius]
 Identities = 48/87 (55%), Positives = 57/87 (65%)

Query: 1  MNMSNINSTRKSSYSHLSATERGRIAAYLKMGKKPVEIARLLGSHRSTICREIKRGSVDQ 60
          MNMS   ST   SY HLS  ERGRI AYL +G KP EIAR LG +RSTI REI RGS+ Q
Sbjct: 1  MNMSTNYSTTNQSYKHLSHAERGRIEAYLSVGLKPAEIARRLGRNRSTITTRRINRGSITQ 60

Query: 61 VKDKNGKQTFFNAYFADSEQRVYETNR 87
          VK  NG++ ++  Y+AD+    Y   R
Sbjct: 61 VKKVNGQKVYYQHYYADAHNRYRHAR 87
```